PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
C12N 15/56, 15/85, C12Q 1/68

A1

(11) International Publication Number: **WO 98/49320**

(43) International Publication Date: 5 November 1998 (05.11.98)

(21) International Application Number: PCT/CA98/00157

(22) International Filing Date: 26 February 1998 (26.02.98)

(30) Priority Data:
2,203,613 24 April 1997 (24.04.97) CA

(71) Applicants *(for all designated States except US)*: HER MAJESTY IN RIGHT OF CANADA, as represented by THE MINISTER OF AGRIGULTURE AND AGRI–FOOD CANADA Central Experimental Farm [CA/CA]; Ottawa, Ontario K1A 0C6 (CA). UNIVERSITY OF OTTAWA [CA/CA]; Research Services, 550 Cumberland, Ottawa, Ontario K1N 6N5 (CA).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: BENKEL, Bernhard, F. [CA/CA]; 293 Summit Avenue, Ottawa, Ontario K1H 5Z7 (CA). NGUYEN, Thuy, H. [CA/CA]; Apartment 1201, 1316 Carling Avenue, Ottawa, Ontario K1Z 7L1 (CA). AHLUWALIA, Navneet, K. [CA/CA]; Apartment 901, 1316 Carling Avenue, Ottawa, Ontario K1Z 7L1 (CA). FALCONER, Marcia, M. [CA/CA]; 33 Abingdon Drive, Nepean, Ontario K2H 7M5 (CA). GAVORA, Jan, S. [CA/CA]; 51 Forest Hill Avenue, Ottawa, Ontario K2C 1P7 (CA). HICKEY, Donal, A. [CA/CA]; 36 Burnham Road, Ottawa, Ontario K1S 0J8 (CA). BENKEL, Kaarina, I. [CA/CA]; 293 Summit Avenue, Ottawa, Ontario K1H 5Z7 (CA).

(74) Agents: ERRATT, Judy et al.; Gowling, Strathy & Henderson, Suite 2600, 160 Elgin Street, Ottawa, Ontario K1P 1C3 (CA).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, GW, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

Published
With international search report.

(54) Title: SECRETED α–AMYLASE AS A REPORTER GENE

(57) Abstract

The regulation of gene expression is of fundamental importance to all biological functions including adaptation to environmental conditions, cell division and differentiation, and the development of disease states such as cancer. The application of "reporter" gene systems, which allow changes in gene expression to be assayed quickly and easily, has contributed greatly to our understanding of the mechanisms involved in gene regulation. A reporter gene system comprising a gene encoding an α–amylase enzyme for use in transforming mammalian cells is described. In a further example the reporter gene system comprises a second α–amylase gene, which codes for an enzyme with an electrophoretically distinct mobility from that of the first α–amylase. The two genes are used for simultaneous expression in mammalian cells, with one isozyme acting as an internal control, and the other isozyme acting as the indicator gene. The use of α–amylase as a secreted reporter gene is particularly attractive because the enzyme is extremely stable over a wide range of conditions. The measurement of α–amylase activity is simple, quantitative, sensitive, safe and inexpensive, and the range of electrophoretically distinguishable variants available allows the assays to be performed in virtually any host without interference from endogenous amylase activities.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	**ES**	Spain	**LS**	Lesotho	**SI**	Slovenia
AM	Armenia	**FI**	Finland	**LT**	Lithuania	**SK**	Slovakia
AT	Austria	**FR**	France	**LU**	Luxembourg	**SN**	Senegal
AU	Australia	**GA**	Gabon	**LV**	Latvia	**SZ**	Swaziland
AZ	Azerbaijan	**GB**	United Kingdom	**MC**	Monaco	**TD**	Chad
BA	Bosnia and Herzegovina	**GE**	Georgia	**MD**	Republic of Moldova	**TG**	Togo
BB	Barbados	**GH**	Ghana	**MG**	Madagascar	**TJ**	Tajikistan
BE	Belgium	**GN**	Guinea	**MK**	The former Yugoslav Republic of Macedonia	**TM**	Turkmenistan
BF	Burkina Faso	**GR**	Greece			**TR**	Turkey
BG	Bulgaria	**HU**	Hungary	**ML**	Mali	**TT**	Trinidad and Tobago
BJ	Benin	**IE**	Ireland	**MN**	Mongolia	**UA**	Ukraine
BR	Brazil	**IL**	Israel	**MR**	Mauritania	**UG**	Uganda
BY	Belarus	**IS**	Iceland	**MW**	Malawi	**US**	United States of America
CA	Canada	**IT**	Italy	**MX**	Mexico	**UZ**	Uzbekistan
CF	Central African Republic	**JP**	Japan	**NE**	Niger	**VN**	Viet Nam
CG	Congo	**KE**	Kenya	**NL**	Netherlands	**YU**	Yugoslavia
CH	Switzerland	**KG**	Kyrgyzstan	**NO**	Norway	**ZW**	Zimbabwe
CI	Côte d'Ivoire	**KP**	Democratic People's Republic of Korea	**NZ**	New Zealand		
CM	Cameroon			**PL**	Poland		
CN	China	**KR**	Republic of Korea	**PT**	Portugal		
CU	Cuba	**KZ**	Kazakstan	**RO**	Romania		
CZ	Czech Republic	**LC**	Saint Lucia	**RU**	Russian Federation		
DE	Germany	**LI**	Liechtenstein	**SD**	Sudan		
DK	Denmark	**LK**	Sri Lanka	**SE**	Sweden		
EE	Estonia	**LR**	Liberia	**SG**	Singapore		

SECRETED α-AMYLASE AS A REPORTER GENE

The present invention relates to a secreted reporter gene system. More specifically, the present invention relates to a secreted reporter gene system based on one or more α-amylases, which are electrophoretically distinct variants of each other, for simultaneous expression in mammalian cells.

BACKGROUND OF THE INVENTION

Reporter genes come in two basic varieties; those that encode intracellular enzymes, and those that code for secreted reporter proteins. Most reporter genes currently in use are of the intracellular variety that require manipulation of the cells, normally harvesting and lysis, prior to measurement of the reporter activity. In contrast, cells transformed with constructs that encode secreted reporter enzymes do not require manipulation or lysis of the cells for activity measurements. Instead, aliquots of the extracellular medium are harvested and assayed directly for enzyme activity. Thus, secreted reporter genes can be used to monitor gene expression *in vitro*, *in vivo* and *ex vivo*.

There are a number of commercially available reporter genes that encode intracellular enzymes, e.g. green fluorescent protein (GFP; Chalfie *et al.*, 1994, Science 263: 802-805), luciferase (de Wet *et al.*, 1987, Mol. Cell. Biol. 7: 725-737), chloramphenicol acetyltransferase (CAT; Gorman *et al.*, 1982, Mol. Cell. Biol. 2: 1044-1051). However, none of these intracellular enzymes are suitable as a secreted reporter activity. For instance, the luciferase reporter gene is used widely due mainly to the sensitivity of the detection system and the lack of background activity in most biological systems. However, the luciferase enzyme is relatively unstable under any conditions, and is completely and irreversibly inactivated by the process of secretion.

Secretable human placental alkaline phosphatase (SEAP); a truncated form of membrane-bound alkaline phosphatase; (Berger, J., Hauber, J., Hauber, R., Geiger, R., Cullen, B.R. 1988. Gene. 66: 1-10; Cullen, B.R., and Malim, M.H. 1992.

Methods Enzymol. 216: 362-368) is a reporter gene that is distributed by Clontech as a secreted reporter gene/enzyme system (see European Patent application 0 327 960). However, many biological systems display endogenous (background) alkaline phosphatase (AP) activity. For example, the high level of alkaline phosphatase activity in cows' milk precludes the use of SEAP as a reporter activity for experiments aimed at using the bovine udder as a "bioreactor" in the production of bioengineered milk. A second commercially-available, secreted reporter gene is human growth hormone (hGH, Boehringer Mannheim; Selden, R.F., Howie, K., Rowe, M.E., Goodman, H., Moore, D. 1986, Mol. Cell. Biol. 6: 3173-3179). However, since hGH has no easily measurable biological activity, the detection system for this secreted product uses a laborious, indirect antibody-based ELISA method.

Thus, there is a need for an improved secreted reporter gene that overcomes the problems of the prior art systems. Specifically there is a need for a secreted reporter gene encoding a biological activity that, unlike hGH, is safe and easily measured, but that does not have the background activity problems associated with SEAP.

SUMMARY OF THE INVENTION

Thus according to the present invention there is provided a secreted reporter gene system for mammalian cells. More specifically the present invention relates to a secreted reporter gene system based on one or more variants of α-amylase, that have different electrophoretic mobilities from that of mammalian α-amylases, and that can be efficiently expressed and secreted by mammalian cells.

The reporter gene system of the present invention comprises a DNA sequence with the following components: 1) a signal peptide coding region; 2) a sequence encoding one or more variants of the α-amylase mature protein; and 3) a transcription termination region.

In one embodiment of the present invention, the α-amylase gene is the chicken (*Gallus gallus*) α-amylase gene.

In one embodiment of the present invention, the α-amylase gene is the fruit fly *(Drosophila melanogaster)* α-amylase gene.

In one embodiment of the present invention, the reporter gene system comprises a DNA sequence based on the α-amylase gene from the chicken (*Gallus gallus*), in conjunction with the α-amylase gene from the fruit fly (*Drosophila melanogaster*), in a dual-reporter plasmid.

In a further embodiment of the present invention, the signal sequence is based on the *Drosophila melanogaster* α-amylase secretion signal.

In yet a further embodiment of the present invention the transcription termination region is based on the *Drosophila melanogaster* α-amylase gene termination region.

In a further aspect of the present invention the reporter gene system also includes a transcription promoter region.

The present invention is also directed to a vector comprising the reporter gene system.

The present invention also defines a procedure to assay for α-amylase activity in complex biological mixtures of proteins.

In yet a further embodiment of the present invention, there is provided a plate diffusion assay, which is used to measure α-amylase activity from a large number of samples.

Also included within the scope of the present invention is a kit including the reporter gene system of the present invention and reagents for testing α-amylase activity.

BRIEF DESCRIPTION OF THE DRAWINGS

These and other features of the invention will become more apparent from the following description in which reference is made to the appended drawings wherein:

FIGURE 1 shows the sequence of the chicken α-amylase cDNA (SEQID No: 1).

FIGURE 2 shows an example of the reporter gene construct which does not contain a promoter region. The α-amylase mature peptide coding sequence '*amylase*', and transcription termination region '*ter*' are indicated. The solid black rectangle between the promoter and the amylase coding region represents the secretion signal peptide. Restriction enzyme recognition sites are as follows: *As* = *Asc*I ; *Hi* = *Hin*dIII; *Nc* = *Nco*I; *Nh* = *Nhe*I; *No* = *Not*I. The 8-cutter sites *Not*I and *Asc*I were built into the upstream and downstream ends. respectively, of the reporter gene in order to facilitate its insertion into the reporter construct.

FIGURE 3 shows an example of the reporter gene construct which contains a promoter region. The description is as above for Figure 2, except the transcription promoter '*pro*' is also indicated.

FIGURE 4 shows the sequence of the transcription termination region (SEQID No: 2).

FIGURE 5 shows the electrophoretic mobilities of α-amylases from 3 different sources in a native gel assay. Lane 1, pig pancreatic amylase; Lane 2, *D. melanogaster* amylase; Lane 3, chicken pancreatic amylase.

FIGURE 6 shows the production of the reporter activity by mammalian cells in culture. Lane 1, chicken pancreatic amylase control; Lane 2, mouse (PG13) cells; Lane 3, mouse (PA317) cells; Lane 4, hamster (CHO) cells; Lane 5, monkey (MA104) cells, Lane 6, cow (ET2) cells; Lane 7, human (Hela) cells; Lane 8, untransformed mouse cells (PG13).

FIGURE 7 shows the expression levels of amylase under different promoters. Lane 1, medium from control, untransformed PA317 cells; Lane 2, cells transformed with reporter construct containing *LTR:amy* reporter gene; Lane 3, cells transformed with reporter contruct containing *Actin:amy* reporter gene; Lane 4, cells transformed with reporter construct containing *CMV:amy* reporter gene; Lane 5, chicken amylase standard.

FIGURE 8 shows the detection of amylase activity in cows' milk. Lane 1, whole milk; Lane 2, whole milk spiked with chicken amylase; Lane 3, empty; Lane 4, chicken amylase standard; Lane 5 & 6, milk following removal of caseins; Lanes 7 & 8, milk with amylase spike following removal of caseins.

FIGURE 9 shows the isolation and detection of α-amylase in cows' pre-lactation secretions (pre-milk). Lanes 1, 2 & 3, early premilk from 3 quarters of cow #99. Lane 1, control quarter (RF-); Lane 2, treated quarter (RH+); Lane 3, treated quarter (LH+). Lanes 4, 5 & 6, late premilk from cow #99. Lane 4, treated quarter (LH+), Lane 5, treated quarter (RH+); Lane 6, control quarter (RF-). Lane 7, mixture of pig and chicken amylase standards; Lane 8, chicken amylase standard.

FIGURE 10 is a schematic depiction of the plasmids used in the dual-reporter system. "Fly" refers to *D. melanogaster*; "chick" refers to *G. gallus*. Abreviations are: ss=signal sequence; amy=α-amylase; CMVp=immediate early promoter of the human cytomegalovirus; LpA=late polyadenylation site; ori=origin of replication; TB=transcriptional blocker; actin p=rat β-actin promoter.

FIGURE 11 shows the production of α-amylase reporter activities by cultured mammalian cells transfected with either the pUC-TEST plasmid or the pUC-CONTROL plasmid. Lane 1, chicken amylase standard; Lane 2, fly amylase standard; Lane 3, medium from control, untransformed PA317 cells, Lane 4, cells transformed with pUC-TEST=base dual reporter construct; Lane 5, cells transformed with pUC-CONTROL=dual reporter construct containing the actin promoter inserted upstream of the fly amylase reporter gene.

FIGURE 12 shows the reporter activites of PA317 cells co-transformed with chicken α-amylase and the firefly luciferase. The reporter genes in each co-transformation were placed under the same promoter: the murine retroviral LTR, the rat β-actin promoter, or the human CMV early immediate promoter. Results for amylase activity were obtained by densitometric scanning of the photograph of the native gel containing the amylase activity obtained from the cell supernatant. Luciferase activity was assayed from cell lysates according to manufacturer's instructions (Promega Corp).

DESCRIPTION OF PREFERRED EMBODIMENT

The present invention relates to a secreted reporter gene system. More specifically, the present invention relates to a secreted reporter gene system based on one or more α-amylases, which are electrophoretically distinct variants of each other, and also distinct from mammalian α-amylases, for simultaneous expression in mammalian cells.

An α-amylase reporter gene encodes a biological activity that is easily measured by a variety of liquid or semi-solid phase assay systems. The use of a native gel assay system eliminates any potential interference by endogenous amylase activities produced by the transformed, mammalian target cells. In addition, the advantage of using a secreted reporter activity is that it is not necessary to harvest cells from the culture dish, or to biopsy a transformed tissue (e.g. udder epithelium), nor is it necessary to lyse cells prior to assay. Instead, the cell culture supernatant or biological fluid (e.g. milk) is recovered from the transformed culture or animal and assayed for reporter activity. This feature has several advantages such as the ability to conduct multiple measurements over time on a single transformed population of cells (time course measurements).

Since α-amylase exists in a number of distinct isoforms, an additional advantage to using the amylase reporter system is that a different isozyme can be used as an internal control in transfection experiments. The "normalizing" isozyme reflects a standard level of gene expression, while the "test" isozyme is the indicator enzyme. Since the isozymes have different migration patterns on the native gel, they can be assayed simultaneously within the same gel lane. This simplifies the assay procedure, in contrast to other reporter systems where the test and normalizing activities require separate assays.

The chicken and *Drosophila* α-amylase enzymes are two examples of suitable α-amylases, which can be used alone or together in both embodiments of the present invention. These enzymes are compatible with the mammalian cellular translation and

secretion machinery. Therefore, the mammalian cells are able to produce and secrete them efficiently in an authentic, bioactive form. Furthermore, these α-amylases have the same physio-chemical requirements for optimal activity and stability, for example, pH optimum, calcium activation, etc., as mammalian amylases. These features are essential for their use as reporter activities. In addition, the electrophoretic mobilities of the mature, active chicken and *Drosophila* α-amylase enzymes on native polyacrylamide gels are different from that of mammalian amylases. Therefore if any amylase enzyme is secreted from mammalian cells, the reporter activities can easily be differentiated from the background activity.

The absence of background activity is particularly important when measuring reporter activities in biological fluids, such as cows' milk, which contains high background levels for other reporter systems, e.g. the SEAP reporter system. According to the present invention, the α-amylase reporter system is most useful in complex protein mixtures such as biological fluids, specifically fetal calf serum, milk, and pre-lactation secretions, because the reporter α-amylase activity can be quantitatively recovered, using the isolation procedure included in this invention, from such fluids, which will be described in detail below.

Included within the scope of the present invention are modified DNA sequences encoding functionally active α-amylases from chicken and *Drosophila*. Possible modifications include but are not limited to: 1) nucleotide substitutions that eliminate restriction enzyme (RE) sites in the naturally-occurring sequence, but do not alter the amino acid sequences, and the use of these RE sites in a multiple-cloning-region for the introduction of DNA components, e.g. the transcription promoter; 2) conservative amino acid substitutions; 3) the enhancement of the specific activity of the reporter polypeptide by amino acid substitutions; 4) increasing the temperature stability of the reporter enzyme by amino acid substitutions; 5) creation of hybrid reporter enzymes from portions of genes coding for different isozymes. Also included is the potential use of amylase inhibitors to specifically inhibit any endogenous, mammalian amylases that might be present in the extracellular medium of the transformed cells.

The reporter gene system of the present invention also includes a DNA sequence encoding a signal peptide coding region, which in one embodiment is based on the *Drosophila melanogaster* α-amylase pre-protein. This signal peptide works effectively in combination with the chicken α-amylase enzyme and the mammalian cellular secretion machinery. The DNA sequence encoding the signal peptide was modified to include useful restriction enzyme sites for the introduction of promoter sequences to drive the reporter gene. The present invention also encompasses any modifications of the signal peptide region that may increase the efficiency of the secretion process beyond the current level.

The reporter gene system of the present invention also includes a transcription termination region. In one embodiment of the present invention the transcription terminator region is derived from the *Drosophila melanogaster* amylase gene. This portion of the reporter gene serves the following functions: (1) to provide a 3' non-coding region for the transcript; (2) to provide a transcription termination signal for the RNA polymerase complex; and (3) to provide a poly-adenylation motif to the primary transcript. In the example where two amylase genes are used in the reporter gene system and one of said genes is the *D. melanogaster* α-amylase gene, the terminator region comes from the late polyadenylation site of SV40, to minimize possible recombination with the α-amylase cassette.

Included within the scope of the present invention is the addition of any functional DNA elements, such as transcription enhancer elements, introns, etc., to the transcription terminator region in order to enhance the overall level of expression of the reporter gene.

In one embodiment of the present invention, the reporter gene system also includes a DNA sequence encoding a transcription promoter region. When the reporter gene system of the present invention includes a promoter, the primary use of this secreted reporter gene system is to monitor the efficiency of the introduction of foreign DNA into mammalian cells *in vivo, in vitro* and *ex vivo*. This application

facilitates experiments designed to optimize the conditions, methods and vehicles employed for the introduction of foreign DNA into mammalian cells. There are a number of high expression level, commercially available promoters which are suitable for this application. Not to be construed as limiting, possible examples include: a Cytomegalovirus promoter, a murine retroviral promoter and a rat beta-actin promoter.

One aspect of the invention would be the use of a tissue-specific promoter to drive the expression of the reporter gene. This configuration would enable the researcher to distinguish between expression in target and non-target cell populations. Furthermore, it would allow researchers to optimize their delivery systems for targeting the foreign DNA to cell types that have the greatest capacity for producing and secreting foreign peptides. This aspect is especially important for *in vivo* gene transfer. Specific applications would include gene transfer to udder epithelial cells in mammals for the production of pharmaceutical agents into the milk (the somatic bioreactor concept), and DNA transfer for gene therapy in humans.

The construction of the reporter gene involves the use of standard recombinant DNA methods, including the use of restriction enzymes to fuse DNA molecules with coherent ends, the amplification of DNA fragments using the polymerase chain reaction (PCR; Canadian patent 1,237,685), and the synthesis of DNA fragments using phosphoramidite chemistry (Matteucci and Caruthers, 1981, J. Am. Chem. Soc. 103: 3185-3191). The application of the reporter gene involves the insertion of the gene into a suitable vector to form double-stranded, circular DNA molecules that are delivered to the eukaryote cells. General methods for the preparation and modification of recombinant DNA molecules have been described by Cohen *et al.* [U.S. patent No. 4,237,224], Collins *et al.* [U.S. patent No. 4,304,863], Sambrook *et al.* (1989, *Molecular cloning: A laboratory manual*, Cold Spring Harbor Laboratory), and Mullis and Faloona (1987, Methods Enzymol. 155: 335-350) . Because they illustrate much of the state of the art, these references are hereby incorporated by reference.

Once assembled, the vector carrying the reporter gene is delivered to the target cells by one of a number of commonly used methods, e.g. viral vectors, calcium phosphate co-precipitation, or liposome-mediated delivery. The DNA plasmid vector used to carry the reporter gene can provide for either transient or stable maintenance of the reporter gene in the target, host cells depending upon the type of vector employed.

The production and secretion of proteins, including reporter enzymes, by eukaryotic cells is a multi-stage process. Briefly, the process begins with the transcription of the gene encoding the secreted product, followed by the processing of the primary transcript into mRNA, and the transport of the mature messenger into the cytoplasm. The mRNA is translated into the preprotein and simultaneously translocated into the first compartment of the secretory pathway by ribosomes associated with the endoplasmic reticulum. Targeting of the preprotein to the secretion pathway is mediated by a secretion signal peptide. Within the endoplasmic reticulum, golgi apparatus and the secretory vesicles, the primary translation product undergoes sorting and maturation to produce the final bioactive protein which is released into the extracellular medium. The present invention provides for novel reporter activities based on the chicken α-amylase gene and other electrophoretically distinct variants of α-amylase, which are compatible with and can be secreted efficiently by mammalian cells. The present invention also includes a secretion signal sequence that effectively targets the reporter peptide to the secretory machinery of mammalian producer cells.

There are a number of prior reports on the expression and secretion of α-amylase in microbial systems. In addition, Japanese Abstract 63263086 described the use of α-amylase as an indicator gene in a "promoter trap" construct for use in both prokaryote and eukaryote systems. The present system however is preferred, in that this system is designed to measure gene expression and DNA delivery specifically in mammalian cells using electrophoretically distinguishable variants in combination with an assay system optimized for use with biological fluids.

The present invention also includes an isolation procedure whereby the reporter activity can be quantitatively recovered from complex protein mixtures such as biological fluids, specifically fetal calf serum, milk, and pre-lactation secretions. The isolation protocol, though based on the procedures of Vretblad, P., 1974, FEBS Lett. 47(1): 86-89, Silvanovich, M. P., & Hill, A. D., 1976, Anal. Biochem. 73: 430-433, and Iefuji, H., et al., 1996, Biochem. J. 318: 989-996, was adapted especially for use in concentrated protein mixtures. Due to the high protein content, the reporter activity cannot be assayed using the original protocols. The present protocol differs from the original procedures in that 1) the non-specific proteins in the fluid are removed by addition of 40% ethanol, and 2) the ethanol-containing solution is applied directly onto the immobilized β-cyclodextrin to further reduce the nonspecific background.

The present invention also includes a "plate diffusion" assay which is one method for measuring the reporter activity when large numbers of samples, for example, culture supernatants from a population of stably selected clones, must be assayed. The method described here is adapted from the observation that colonies of amylase-producing microorganisms form clearing-zones on starch plates (Iefuji, H., et al. 1994. Biosci. Biotechnol. Biochem. 58: 2261-2262). In the amylase reporter system, the amylase-containing solutions are applied to wells formed in starch-containing agarose plates; diffusion of the solution through the starch-agarose forms clearing zones whose size correlates with the concentration of amylase in the solution.

Under conditions where cells are transiently transfected with the reporter plasmid, the amount of secreted reporter activity measured in the extracellular medium of a population of cells, transformed with a multicopy plasmid carrying a reporter gene which is driven by a strong, systemic promoter, is determined by two variables: 1) the number of cells that have received the reporter construct and; 2) the average number of copies of the DNA construct that each transformed cell received. Thus, under these conditions, the secreted reporter activity provides a measure of the efficiency of the DNA delivery system, i.e. how many cells in the population were transformed with

foreign DNA and how many copies, on average, of the transforming vector were incorporated into each transformed cell.

Under conditions where the cells are stably transfected with the reporter plasmid, the level of secreted reporter activity would, in addition, depend on: 1) the site at which the reporter gene becomes integrated, 2) the metabolic state of the cells, and 3) the environmental conditions that may regulate the cell's metabolism in general, and the promoter that drives the reporter gene in particular. If two of the above conditions are experimentally controlled, the secreted reporter gene can be used to monitor the effects of the third variable in time-course-type experiments.

The present invention also includes amylase reporter kits. Kits containing a single or double amylase gene are included within the scope of the present invention.

The single reporter kit would include: two DNA constructs; a plate assay means; and instructions for use. The single reporter kit would optionally include a beta-dextrin purification system, which would be needed in high-protein fluids to allow detection of the amylase activity.

The first DNA construct includes a test plasmid containing a single gene coding for a single amylase isozyme, with a multiple cloning region just upstream where a test promoter can be inserted. The second DNA construct includes a control plasmid where the gene is placed under control of a strong promoter. The plate assay detects amylase activity very simply. It consists of : a mould to form the gel plate; a buffer used to make the agarose media; detection solutions; and a sample of amylase as a positive control. The beta-dextrin system, if included, removes non-specific proteins from high-protein fluids to allow detection of the amylase activity. It consists of: an immobilized beta-dextrin packed into columns; equilibration buffer; wash buffer; and elution buffer. The buffers could be included within the kit or be substituted with instructions on the buffer preparation. The kit will also include a protocol booklet, which would give detailed procedures, as well as simple steps used

to remove background signals from high-protein biological fluids such as serum-containing media, pre-lactation secretion, and milk.

The dual amylase report kit will include: two DNA constructs; a gel electrophoresis assay means; and instructions for use. The dual reporter kit would optionally include a beta-dextrin purification system, which would be needed in high-protein fluids to allow detection of the amylase activity.

In this system the first DNA construct includes a test plasmid (containing the genes coding for both amylase isozymes, one of which is under a "normalizing" promoter). The second DNA construct includes a control plasmid containing the isozymes under different promoters. Gel electrophoresis allows unambiguous identification of the 2 isozymes. The assay consists of: a sample buffer; gel preparation buffer; detection solutions and amylase isozymes as positive controls. As with the single amylase reporter kit the buffers can be included within the kit or they can be replaced with instructions on the buffer preparation.

If included, the beta-dextrin system is identical to that contained in the Single Reporter Kit and the protocol booklet is identical to that contained in the Single Reporter Kit.

The kit of the present invention is much more inexpensive than any expression kit presently available. Furthermore, amylase is a very stable enzyme, whose signal improves with duration of expression. The dual expression kit allows for normalization of the expression data. Endogenous levels, as mentioned previously are not a problem with this detection system.

While this invention is described in detail with particular reference to preferred embodiments thereof, said embodiments are offered to illustrate but do not limit the invention.

EXAMPLES

EXAMPLE 1: Chicken α-amylase cDNA

In order to isolate a chicken genomic α-amylase clone, a commercial chicken library in the vector λ FIX II (Stratagene) was screened with a murine amylase cDNA probe (1.2 kbp *Pst*I fragment of pMPa21, Hagenbuechle *et al.*, 1980, Cell 21: 179-187). Hybridization was carried out under conditions of reduced stringency at 37°C in a standard buffer system containing 50% formamide (Benkel and Gavora, 1993, Animal Genetics 24: 409-413). Following hybridization, the filters were washed twice for 15 min each in 2x SSC, 0.1% SDS at 42°C, followed by a single wash in 0.5x SSC, 0.1% SDS at 50°C for 30 min. Autoradiograms were exposed at -70 C with intensifying screens.

Eighteen positive signals were detected following the first round of library screening. Ten plaques were chosen at random for second screening, and 4 of these isolates were still positive after a third round of screening. One clone (λA-1) was chosen for large-scale DNA isolation and fragment subcloning in preparation for sequence analysis.

A series of overlapping subclones spanning the amylase genomic region was constructed by inserting restriction fragments of the primary isolate λA-1 into the vector pUC18. Double-stranded DNA was sequenced by the gene-walking method using synthetic oligonucleotide primers. Primers were synthesized on an Applied Biosystems 392 synthesizer and deblocked and desalted before use. Sequencing reactions were performed using the dye terminator cycle sequencing kit as described in the instructions supplied by the manufacturer (Applied Biosystems Inc.). The extension products were analyzed on an Applied Biosystems 373A automated sequencer, and sequence assembly was performed using MicroGenie software by Beckman.

The chicken amylase coding region was prepared using reverse transcription-polymerase chain reaction technology (RT-PCR). Oligonucleotide primers for the PCR

step were designed based on the chicken genomic amylase sequence derived as described above. Approximately 2 μg of total RNA from chicken pancreas was used as substrate in a reverse transcription (RT) reaction using the Perkin Elmer RT-PCR kit components according to the instructions supplied by the manufacturer. The RT-reaction was primed with oligo-dT. The oligonucleotide primers used for the PCR stage of the RT-PCR reaction were as follows: (i) *Chamy-Nhe* (5'-ATGCTAGCTCAGTACAATCCCAACACTCAGGCT-3', SEQID No: 3) which spans the position in the coding region corresponding to the N-terminus of the mature enzyme; and (ii) *Chamy-Hin* (5'-CGAAGCTTATAACTTGGCATCA ACGTGAATTG-3', SEQID No: 4) which spans the stop codon of the amylase coding region. The primers were designed to amplify the region in the gene encoding the mature α-amylase peptide. In addition, *Chamy-Nhe* converts the environment of the signal peptidase cleavage site into an *Nhe*I site, while *Chamy-Hin* adds a *Hin*dIII site immediately downstream of the stop codon (enzyme recognition sites are underlined). The modifications introduced by the PCR primers allow the amplified cDNA to be inserted into the reporter construct using cohesive overhang ligation.

PCR amplification of the chicken amylase cDNA was performed using the LA-PCR kit as described in the manufacturer's product bulletin (TaKaRa). The amplified fragment was sequenced using PCR-based chain termination technology, prior to insertion into the reporter construct - this sequence, SEQ ID No: 1, is shown in Figure 1.

EXAMPLE 2: Signal Peptide Coding Region

The signal peptide used in the current reporter gene is modelled on the sequence found in the *Amy*1 gene of the Oregon-R strain of *Drosophila melanogaster* (Boer and Hickey, 1986, Nuc. Acids Res. 14: 8399-8411). The sequence of the native signal peptide extends from nucleotide 1 to 54 in Figure 2 of that reference. The 5'-end of the sequence encoding the signal peptide was modified to incorporate an *Nco*I site straddling the translation start codon (ATG) - i.e. the sequence AT ATG T was changed to CC ATG G. This modification changes the amino acid immediately downstream of the initiator

Met from a Phe to a Val. In addition, the 3'-end of the signal sequence was modified to accommodate an *Nhe*I site. Here, the original *D. melanogaster* sequence of --Ala-Asn-Ala-- (GCC AAC GCC) was changed to --Ala-Leu-Ala-- (GCG CTA GCC). The DNA fragment encoding the signal peptide was constructed in the laboratory on a DNA synthesizer, and assembled into the reporter construct using the restriction enzyme sites built into its ends. The novel signal peptide sequence used in the reporter construct has the following sequence (SEQID No: 5).

```
5'-CCATGGTTCTGGCCAAGAGCATAGTGTGCCTCGCCCTCCTGGCGGTGGCGCTAGCT-3'
3'-GGTACCAAGACCGGTTCTCGTATCACACGGAGCGGGAGGACCGCCACCGCGATCGA-3'
   NcoI                                                  NheI
```

EXAMPLE 3: Promoter Region

In this example the immediate early promoter of the Human cytomegalovirus (CMV; Stinski and Roehr, 1985, J. Virology 55: 431-441) was used to drive the expression of the reporter gene.

Other examples of promoters that have been used successfully in the reporter gene include; a murine retroviral LTR, which is a composite of the MMTV (Mouse Mammary Tumor Virus; Ponta *et al.* 1985, Proc. Natl. Acad. Sci. 82: 1020-1024) and the MoMSV (Moloney Murine Sarcoma Virus; Lin *et al.* 1990, Proc. Natl. Acad. Sci. 87: 36-40) LTRs, and the rat beta-actin promoter (Nudel et al. 1983, Nucl. Acids Res., 11, 1759-1771).

DNA fragments to be used as promoters were amplified by PCR using oligonucleotide primers that incorporated useful restriction enzyme sites at the upstream (*Not*I) and downstream (*Nco*I) ends of the amplified promoter sequences. Restricted promoters fragments were inserted into the reporter gene construct by cohesive end ligation.

EXAMPLE 4: Transcription Termination Region

The final component of the reporter gene is the transcription terminator. In this example 567 bp of the *Drosophila melanogaster* Oregon-R amylase gene termination region was used. The sequence of this section (SEQID No: 2) is shown in Figure 4. The DNA fragment was amplified by PCR using primers that incorporated useful restriction enzyme sites. A *Hin*dIII site was added just upstream of the TAA (stop codon) of the amylase coding region. This TAA is out of frame with the chicken amylase coding region, which brings in its own stop codon. The downstream end of the terminator is modified into an *Asc*I site.

EXAMPLE 5: *Drosophila melanogaster* α-amylase cDNA

Plaque lifts of a lambda gt10 library containing cDNAs from RNA of a *Drosophila melanogaster* Oregon-R strain were hybridized to a mouse salivary α-amylase cDNA probe (the 1650 bp *Pst*I restriction fragment of pMSa104; Hagenbuechle *et al.*, 1980, Cell 21: 179-187), under conditions of reduced stringency - 38°C in a hybridization buffer containing 50% formamide. Following hybridization, filters were washed in 1 X SSC, 0.1% SDS at 52°C.

Eight postive plaques were identified from an estimated 20,000 plaques screened. The longest of these cDNA inserts (OR-M7) was subcloned into a plasmid vector (pUC13) and subjected to DNA sequencing by standard methods (Benkel et al., Genome 29: 510-515). The DNA sequence was determined as described in Boer & Hickey, 1986, Nucl. Acids Res. 14: 8399-8411). The extension products were analyzed on an Applied Biosystems 373A automated sequencer, and sequence assembly was performed using MicroGenie software by Beckman. OR-M7 was found to contain the complete coding region of the *Drosophila melanogaster* α-amylase pre-enzyme.

The DNA fragment encoding the mature *Drosophila melanogaster* α-amylase enzyme was inserted into expression vectors by PCR amplification using the LA-PCR kit and synthetic oligonucleotide primers. The oligonucleotide primers used for the PCR amplification were as follows: (i) *Dm-Q1* (5'-ATGCTAGCGCAATTCGACAC

CAACTACGCA-3', SEQID No: 6) which spans the position in the coding region corresponding to the N-terminus of the mature enzyme; and *Dm3-2* (5'-CAAGCTTGTTT ACAACTTGGCGTTGACGT-3', SEQID No. 7) which spans the stop codon of the amylase coding region. The primers were designed to amplify the region in the gene encoding the mature α-amylase peptide. In addition, *Dm-Q1* converts the environment of the signal peptidase cleavage site into an *Nhe*I site, while *Dm3-2* adds a *Hin*dIII site immediately downstream of the stop codon (enzyme recognition sites are underlined). The modifications introduced by the PCR primers allow the amplified cDNA to be inserted into the reporter construct using cohesive overhang ligation. The modified *Drosophila* α-amylase cDNA is SEQID No. 8.

EXAMPLE 6: The α-amylase dual-reporter gene system

Construction of the "test" and "control" plasmids began with the plasmid pIBI-B, which was formed from the vector pIBI25 (International Biotechnologies Inc.) by techniques of PCR amplification using primers with redundant overhangs (Scharf, 1990), and standard oligonucleotide linker addition (Sambrook et al., 1989). pIBI-B consists of: (i) *Drosophila* α-amylase promoter; (ii) the *Drosophila* α-amylase signal sequence, as described in Example 2; and (iii) the *Drosophila* α-amylase termination region, as described in Example 4.

The region spanning the chicken α-amylase cDNA was amplified by long-range PCR (LA-PCR kit, TaKaRa), incorporating a *Nhe*I site at the 5' primer and a *Hind*III site at the 3' primer. This PCR product was ligated into the plasmid pIBI-B via the *Nhe*I and *Hind*III sites to form the plasmid pIBI-fAmy, which contains chicken α-amylase joined to *Drosophila* amylase signal sequence. The CMV promoter was amplified by PCR from the plasmid pLNCX (Miller, D.M., & Rosman, G.J., 1989, BioTechniques, 7(9): 980-990), then inserted into the plasmid pIBI-fAmy, just upstream of the 5' end of the *Drosophila* signal sequence, replacing the *Drosophila* α-amylase promoter and forming the plasmid pIBI-CfAmy. The entire region spanning the CMV promoter to the 3' end of the chicken α-amylase gene was amplified by long-range PCR, incorporating unique sites at the 3' and 5' ends. This PCR

product was ligated to the 3110bp *Hind*III-*Fse*I fragment of the plasmid pSEAP2-Basic (Clontech), forming the plasmid pUC-CFAmyHF. The *Drosophila* α-amylase gene, including the signal sequence, was amplified by PCR from the cDNA clone; the PCR product was inserted upstream of the chicken α-amylase cassette in the plasmid pUC-CFAmyHF to form the final plasmid pUC-TEST (SEQID No: 9). This "test" plasmid, shown in Figure 10, has the chicken α-amylase gene driven by the CMV promoter. The *Drosophila* α-amylase gene is not under any promoter; instead, there is a polylinker sequence 5' of the *Drosophila* α-amylase gene that allows promoters of interest to be inserted. In SEQID No:9 the components are located in the following positions: *Drosophila* α-amylase signal sequence: 83-134; *Drosophila* α-amylase: 135-1575; *Drosophila* α-amylase terminator: 1588-2128; CMV promoter: 2129-2951; *Drosophila* α-amylase signal sequence: 2952-2999; chicken amylase: 3000-4505; SV40 late polyadenylation site: 4506-4754; pUC origin of replication: 5033-5676; β-lactamase gene: 6684-582; β-lactamase promoter region: 6754-6685; F1 origin: 6816-7272; synthetic polyadenylation site: 7462-7450; and pause site from human α2-globin gene: 7464-7555.

To construct the plasmid pUC-CONTROL, the rat β-actin promoter was inserted between the transcription block and the 5' end of the *Drosophila* α-amylase gene in the plasmid pUC-TEST. This plasmid (SEQID No: 10), shown schematically in Figure 10, has the chicken α-amylase gene driven by the CMV promoter, and the *Drosophila* α-amylase gene driven by the β-actin promoter. In SEQID No:10 the components are located in the following positions: pUC origin of replication: 5384-6027; β-lactamase gene: 7035-6175; β-lactamase promoter: 7105-7036; F1 origin: 7167-7623; synthetic polyadenylation site: 7753-7802; and pause site (human α2-globin): 7815-7906.

EXAMPLE 7: The Chicken α-amylase reporter gene system

A variety of methods are available for the delivery of the reporter gene to the target host cells including: 1) liposome-mediated fusion; 2) micro-injection; 3) particle bombardment; 4) viral vectors; and 5) calcium phosphate co-precipitation.

For the purpose of this non-limiting illustrative example, the reporter gene linked to the CMV promoter has been inserted into the plasmid pIBI25 (International Biotechnologies Incorporated) to produce a transient expression construct. In addition, the reporter constructhas been delivered to the cells using the calcium phosphate co-precipitation method. Other methods of DNA delivery are applicable, as is the use of vectors that promote the stable maintenance of the reporter gene within the transformed cells, e.g. retroviral vectors.

EXAMPLE 8: Harvesting of the Reporter Activity from Serum-Free Culture Supernatants:

Cells were grown under standard mammalian cell culture conditions. The media used to culture the cell lines were as follows: (i) for MA104 and Hela, Dulbecco's modified Eagle's medium (DMEM) with 10% Fetal Bovine Serum; (ii) for CHO, DMEM with 10% Newborn Calf Serum; (iii) for PA317 and PG13, DMEM with iron-enriched 10% Calf Serum. Cells were plated at a density of $5x10^5$ cells per 10 cm dish. Calcium phosphate transfection was carried out overnight. Following transfection, cells were incubated overnight in cell line-specific media as described above. On day 4, the culture medium was removed and replaced with Serum-free DMEM for all cell lines. On day 5, samples of medium were collected, and debris was removed from the samples by centrifugation at 500xg.

Cell supernatant was adjusted to 40% (vol/vol) of ethanol and the mixture incubated on ice for 1 hour. The solution was centrifuged at 16,000xg for 20 minutes at 4°C and the supernatant dried under vacuum to remove the ethanol. 100-150μL of the resulting solution was used for routine analysis by native gel electrophoresis. To concentrate the solution, ~1mg of carrier protein (e.g. BSA) was added to the sample and the solution applied to a Centricon-30 filtration columns (Amicon). As little as 0.0005U of amylase processed in this manner can be detected by native gel electrophoresis.

EXAMPLE 9: Harvesting of the Reporter Activity from Serum-Containing Culture Supernatants:

Cells were grown under standard mammalian cell culture conditions as described above, except that 10% serum was included in the culture media. The procedure used to harvest the reporter protein is the same as that described for harvesting from milk (see below). The presence of serum gives some non-specific activity; however, this activity is clearly distinguishable from the reporter activity when detected by native gel electrophoresis (described in Example 12) because it has a different mobility. For an absolutely clean background and for concentrating the sample, adsorption on β-cyclodextrin (described in Example 11) can be used.

EXAMPLE 10: Harvesting of the Reporter Activity from Milk:

Two alternative methods are used to recover the α-amylase activity from milk. In the first method, fresh milk containing α-amylase was de-fatted by centrifugation at 8000xg, 10 min. at room temperature. The aqueous fraction was acid-curdled with 1M HCl to a final pH of 4.5, then centrifuged at 8000xg for 20 min. The whey (supernatant) fraction was neutralized with 1M NaOH to a final pH of 6.7. A portion of the final solution was used for analysis.

In the second method, the de-fatted milk was adjusted to 40% ethanol, placed on ice for 1-2h, then centrifuged at 16000xg at 4°C. The supernatant was then placed under vacuum to remove the ethanol, and a portion of the final solution was used for analysis with essentially complete recovery of the reporter activity. For concentrating the sample, adsorption on β-cyclodextrin without removing the ethanol (described in Example 11) can be used.

EXAMPLE 11: Harvesting of the Reporter Activity from Pre-laction Secretion (Pre-milk):

Bovine pre-milk is a jelly-like substance that is so viscous it is nearly solid. To analyze α-amylase activity in the pre-milk, the sample was diluted 1/3-1/4 with water, then adjusted to 40% ethanol. The sample was placed on ice for 1-2h, then centrifuged

at 16000xg at 4°C to remove the precipitated proteins. β-cyclodextrin-Sepharose 6B, prepared as described in Vretblad, P., 1974, FEBS Lett. 47(1): 86-89, was added to the sample at a ratio of 40μL packed beads to 4mL of prepared sample. The sample was shaken for 2h at room temperature, then washed 5 times in 1.5mL of wash buffer (50mM Tris-HCl, 1mM $CaCl_2$ pH7.5). The sample was eluted by incubation for 5 min at room temperature in a desired volume of 10mg/mL β-cyclodextrin. As little as 0.0002U of α-amylase activity per mL of pre-milk can be recovered this way.

EXAMPLE 12: Detection of Reporter Activity

Amylase activity can be detected and measured by a variety of simple, safe procedures. These can be divided into the tube or liquid assays and the assays that involve diffusion in a semi-solid medium or electrophoretic separation.

The preferred protocol is electrophoretic separation of the proteins in the sample on a native (non-denaturing) polyacrylamide gel ("native gel electrophoresis"). This protocol is described in Benkel and Hickey (1986, Genetics 114: 137-144). Briefly, a low percentage acrylamide gel allows proteins to be separated on the basis of the overall charge of the molecules. Thus, even though chicken and mammalian amylases are very similar in molecular weight, they display very different migration patterns in the native gel assay (see Figure 5).

Following electrophoretic separation, the gel is incubated in a buffer solution containing partially-hydrolyzed starch. The starch granules coat the gel and penetrate the gel surface. Staining of the starch-coated gel with iodine results in a gel that shows clear amylase bands on a dark blue background.

The main advantage of the gel assay is that the background activity measured for the recipient cells is essentially zero. The gel assay is highly sensitive, and can easily be converted into a quantitative format by the incorporation of serial dilutions of an activity standard (see Benkel and Hickey, 1986, Genetics 114: 943-954). In

addition, there is a dye-linked starch-based substrate that facilitates real-time activity visualization.

As an alternative to native gel electrophoresis, prepared samples containing α-amylase can also be assayed by a very simple "plate diffusion" method. 1% agar, made in 50mM Tris-HCl, 1mM $CaCl_2$ pH 7.5, containing 0.02% starch, was poured onto culture plates. 4mm-diameter holes were made in the agar, into which up to 30μL of sample can be placed. The plates were covered and incubated overnight at 37°C to allow the amylase to diffuse. "Haloes", formed as the starch becomes digested, were detected by staining with iodine as described in the native gel electrophoresis procedure.

Chicken amylase in cell culture medium and in milk was tested and the activity was found to be stable for days at room temperature in a bioactive format. In addition, freezing and thawing does not affect the activity of the enzyme.

Figure 6 shows a variety of mammalian cell lines transformed with the reporter construct containing the CMV promoter delivered using a transient-expression vector. All cell lines tested including; CHO (hamster), ET-2 (cow), Hela (human), MA104 (monkey), PA317 and PG13 (mouse) showed production and secretion of chicken amylase according to the native gel assay. Different cell lines appear to secrete chicken amylase at different efficiencies. This may reflect the ability of the cells to take up foreign DNA, or it may be a measure of the inherent capacity of the different cell types to secrete proteins.

Figure 7 shows the results of an experiment in which PA317 cells were co-transfected with two reporter constructs: one containing chicken α-amylase as a reporter; the other contains firefly luciferase as a well-known intracellular reporter. In each co-transfection, both reporters are driven by the same promoter, which is the murine retroviral LTR, the rat β-actin promoter, or the human CMV early immediate promoter (see EXAMPLE 3 above for details of the promoters). The supernatants of

transiently transfected cells were harvested and analyzed for reporter activity as described in Example 8, while cell pellets were processed and analyzed for luciferase activity according to the instructions supplied by the luciferase assay kit manufacturer (Promega Corp). The results for extracellular amylase reporter activities show that the expression level of the CMV promoter is roughly four times that of the β-actin promoter, which in turn is about twice that of the LTR. Similar results were obtained with the intracellular luciferase promoter (Figure 12).

One of the primary applications for secreted reporter genes is in the optimization of DNA delivery and transformation *in vivo* of mammalian cells that secrete proteins in to biological fluids. This optimization process is crucial to the development of successful approaches for cell transformation in gene therapy. We are currently exploring the use of somatic transgenesis of bovine epithelial tissue in the development of 'bioreactor' cows for the production of pharmaceutical agents into milk. In order to compare the suitabilities of SEAP (Clontech Cat. No. K2041-1) and α-amylase as reporter gene/enzyme systems for the transformation of cow udder cells, SEAP and α-amylase were spiked into fresh cows' milk their activities measured. The SEAP used to spike the milk came from cell culture supernatant; the α-amylase was prepared from chicken pancreas.

Fresh cows' milk was centrifuged at 8,000 xg for 10 minutes at room temperature to remove the fat. SEAP cell supernatant and chicken α-amylase was added to the aqueous fraction. For the SEAP reconstitution experiment, SEAP activity was measured for the following treatments: (i) untreated milk aqueous phase ("Endogenous" in Table 1); (ii) milk aqueous phase heated to 65°C for 30 minutes ("Heated"); (iii) milk aqueous phase with SEAP spike ("Spiked"); and (iv) milk aqueous phase with SEAP spike heated to 65°C for 30 minutes ("Spiked/heated").

The results demonstrated that there is a high background of endogenous, soluble alkaline phosphatase activity in cows' milk. This activity is indistinguishable from SEAP. Heating the samples to 65°C does not decrease the endogenous SEAP activity

level; instead it appears that a portion of the endogenous alkaline phosphatase activity is present in an inactive, protein-complexed state in milk, and that heating tends to promote the release of some of this protein-bound endogenous SEAP activity. This results in a further increase in the background SEAP levels measured in the unspiked milk. The ability of milk proteins to inactivate SEAP is also evident from the dramatic decrease in SEAP signal measured in the spiked samples, where only about 1/4 of the SEAP activity added in the spike is detectable following the addition of the milk. Heating the spiked samples to 65°C raises the levels somewhat on average, but this increase in activity is at least partially due to the release of endogenous alkaline phosphatase enzyme.

TABLE 1: Measurement of SEAP activity in cow's milk.

Sample	Endogenous	Heated	Spiked	Spiked/heated
Cow 1	52	60	282	307
Cow 2	59	62	198	190

Legend: Cows' milk was processed as described in the text. SEAP activities are given in 1000's of luminometer units. "Endogenous" refers to the endogenous background activity present in the aqueous phase of the milk; "Heated" refers to the background activity in the milk following a 30 minute incubation at 65°C; "Spiked" refers to the unheated milk following the addition of the SEAP spike; "Spiked/heated" represents the spiked sample following a 30 minute incubation at 65°C. The SEAP spike consisted of 1,075 (x1000) units.

Increasing the level of the alkaline phosphatase-inhibitor homoarginine in the samples had no effect on the SEAP activity measurements. On the other hand, acid curdling of the milk to remove caseins, a standard method used to fractionate milk, resulted in the loss of the bulk of the SEAP spike. Likewise, adjusting the milk to 40% ethanol eliminates the SEAP spike, which precipitated out with the non-specific proteins.

Figure 8 shows milk samples with and without the addition of chicken α-amylase preparations. For whole milk, the gel assay shows a light smear throughout

the entire lane (see lane 1). This is due to the effect of the abundant milk casein proteins. However, the band of activity corresponding to the chicken amylase reporter 'spike' is clearly visible even in whole milk (lane 2). Our standard protocol for the amylase reconstitution experiments included an acid curdling step to remove caseins from the whole milk. This treatment was incompatible with the SEAP reporter system because it almost entirely depleted the milk aqueous phase of alkaline phosphatase activity (see above).

Following neutralization, the whey containing the chicken α-amylase spike was analyzed by the standard gel assay. This treatment completely eliminated any endogenous background activity from the milk samples, and left the chicken amylase band clearly visible in the spiked samples (lanes 7&8). Thus, chicken α-amylase is superior to SEAP as a secreted reporter enzyme in biological fluids such as cow's milk.

As outlined above, a major advantage of the amylase reporter system over existing reporter assays is the ability to detect chicken amylase activity in complex biological fluids. Cow's pre-lactation secretions (also known as pre-milk) are particularly problematic biological samples due to their extremely high content of proteins, fats and carbohydrates. However, using the simple β-cylodextrin adsorption procedure described in Example 11, one can recover the chicken α-amylase reporter activity amylase quantitatively even from such recalcitrant materials such as early and late pre-lactation secretions of pregnant heifers (see Figure 9). Note the complete lack of background activity for chicken amylase in the pre-milk of untreated udder quarters (lanes 1 & 6), and the presence of reporter activity only in udder quarters that were trans-infected with constructs encoding chicken α-amylase (lanes 2, 3, 4 & 5). In this way, as little as 0.0002U of α-amylase activity per mL of pre-milk can be detected.

Figure 10 gives a schematic representation of the dual reporter constructs used for promoter strength determinations in transiently transfected, culture cells. To measure the strength of a "test" promoter, the pUC-TEST plasmid is used. It consists

of two expression cassettes. The *Drosophila* α-amylase cassette has the *Drosophila* α-amylase cDNA, including the signal sequence and the terminator region. There is no promoter driving this reporter gene; instead, a polylinker region is provided immediately 3' of the cDNA to facilitate insertion of a promoter sequence of interest. Just upstream of this polylinker region is a partially synthetic "transcriptional blocker" designed to minimize non-specific transcription of α-amylase. The *Drosophila* α-amylase gene is, therefore, the "test" isozyme. The second cassette has the chicken α-amylase cDNA joined at the 5' end with the signal sequence from *Drosophila melanogaster*; the terminator region for this cassette comes from the late polyadenylation site of SV40, to minimize possible recombination with the *Drosophila* α-amylase cassette. The chicken α-amylase gene is driven by the CMV promoter, and therefore, acts as the "normalizing" isozyme. The backbone of the plasmid pUC-based.

To provide a control showing that both cassettes are functional, the plasmid pUC-CONTROL is used. This plasmid is identical to pUC-TEST, except that the rat β-actin promoter was inserted upstream of the *Drosophila* α-amylase.

Figure 11 shows the results of transfecting PA317 cells in culture with the dual reporter construct. Lane 3 shows the lack of background activity in the culture medium of untransfected PA317 cells. Lane 4 contains culture medium from PA317 cells transfected with the plasmid pUC-TEST. The activity band is specific to the normalizing activity resulting from the expression of the chicken amylase reporter gene under the control of the CMV promoter. Lane 5 contains culture medium from PA317 cells transfected with the vector pUC-CONTROL. The chicken amylase band results from the expression of the normalizing isozyme while the *Drosophila melanogaster* (fly) amylase activity band results from the expression of the fly gene under the control of the β-actin promoter. The relative signal intensities of the two activity bands within the same gel provide a direct comparison of the transcriptional activities of the CMV and β-actin promoters. The β-actin promoter was inserted into pUC-CONTROL for the purposes of this example only. It will be obvious to persons skilled in the art that

the pUC-TEST assay system can be used to assay any nucleotide sequence whether natural or synthetic for promoter activity. Furthermore, by employing host cells other than PA317 cell line, the assay can be used to evaluate the effect of host cell type on promoter activity. In this particular instance, 80μL of culture supernatant, processed as described in Example 8, was used.

All scientific publications and patent documents are incorporated herein by reference.

The present invention has been described with regard to preferred embodiments. However, it will be obvious to persons skilled in the art that a number of variations and modifications can be made without departing from the scope of the invention as described in the following claims.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT:
(A) NAME: Her Majesty in Right of Canada as Rep. by Agriculture and Agri-Food Canada
(B) STREET: Central Experimental Farm
(C) CITY: Ottawa
(D) STATE: Ontario
(E) COUNTRY: Canada
(F) POSTAL CODE (ZIP): K1A 0C6

(A) NAME: The University of Ottawa
(B) STREET: 115 Seraphin Marion
(C) CITY: Ottawa
(D) STATE: Ontario
(E) COUNTRY: Canada
(F) POSTAL CODE (ZIP): K1N 6N5

(ii) TITLE OF INVENTION: Secreted alpha-Amylase as a Reporter Gene

(iii) NUMBER OF SEQUENCES: 10

(iv) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)

(v) CURRENT APPLICATION DATA:
APPLICATION NUMBER: CA 0,000,000
(vi) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: CM 2,203,613
(B) FILING DATE: 24-APR-1997

(2) INFORMATION FOR SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1505 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

GCTAGCTCAG TACAATCCCA ACACTCAGGC TGGGAGGACA TCTATCGTGC ATCTCTTTGA 60

ATGGCGCTGG GCCGACATTG CACTGGAGTG CGAACACTAT TTAGCTCCTA ATGGGTTTGG 120

AGGAGTTCAG GTTTCTCCTC CAAATGAAAA CATTGTCATT ACTAATCCGA ACAGGCCCTG 180

GTGGGAAAGA TACCAGCCCA TCAGCTACAA GATCTGCAGT CGATCGGGCA ATGAAAATGA 240

ATTCAGAGAC ATGGTGACCA GATGCAACAA TGTTGGAGTT CGTATTTATG TGGATGCTGT 300

TGTCAATCAC ATGTGTGGAT CTATGGGTGG CACGGGCACC CACTCAACAT GTGGGAGCTA 360

TTTCAACACC GGGACTAGAG ATTTTCCCGC TGTGCCGTAC TCTGCCTGGG ATTTCAATGA 420

CGGCAAATGT CACACTGCAA GTGGAGACAT CGAAAATTAT GGGGACATGT ATCAGGTCCG 480

GGATTGCAAG TTGTCCAGCC TTCTTGATCT GGCTCTGGAG AAGGACTATG TACGCTCAAC 540

AATTGCAGCG TACATGAATC ACCTCATTGA TATGGGTGTA GCAGGGTTCC GGATCGATGC 600

TGCCAAGCAT ATGTGGCCAG GGGACATAAG AGCGTTTCTG GACAAACTGC ACGATCTAAA 660

TACTCAGTGG TTTTCAGCAG GAACGAAACC CTTTATTTAC CAAGAGGTAA TTGACTTGGG 720

AGGAGAGCCA ATCACAGGCA GTCAGTACTT TGGGAATGGC CGCGTGACAG AATTCAAGTA 780

TGGTGCCAAA CTGGGGACGG TGATCCGGAA GTGGAATGGA GAGAAGATGG CCTACTTAAA 840

GAACTGGGGA GAAGGCTGGG GCTTTGTGCC TTCTGACAGA GCCCTGGTGT TTGTGGATAA 900

CCACGACAAC CAGCGGGGGC ACGGGGCAGG CGGAGCTTCC ATTCTTACTT TCTGGGATGC 960

CAGGCTTTAT AAAATGGCGG TTGGTTTCAT GCTCGCTCAT CCGTACGGGT TCACACGGGT 1020

GATGTCAAGT TATCGTTGGC CAAGATATTT CGAAAACGGA GTGGATGTTA ACGACTGGGT 1080

GGGACCACCA AGTAACTCGG ACGGATCGAC GAAGTCCGTT ACAATCAACG CAGACACTAC 1140

CTGTGGCAAT GACTGGGTCT GCGAACATCG CTGGCGACAA ATAAGGAACA TGGTTATCTT 1200

CCGTAATGTG GTAGACGGTC AGCCTTTCTC AAACTGGTGG GACAACGGGA GCAATCAAGT 1260

AGCTTTCGGT CGCGGCGACA GAGGCTTCAT TGTCTTTAAT AATGATGACT GGTATATGAA 1320

TGTCGATTTG CAAACTGGTC TGCCTGCTGG AACCTACTGC GATGTTATTT CTGGACAAAA 1380

GGAAGGCAGT GCGTGTACTG GAAAGCAGGT GTACGTTTCC TCGGATGGAA AGGCCAATTT 1440

CCAGATTAGT AACAGCGATG AAGATCCATT TGTTGCAATT CACGTTGATG CCAAGTTATA 1500

AGCTT 1505

(2) INFORMATION FOR SEQ ID NO: 2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 568 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

AAGCTTGTAA ACAGCTGGGG AGCATGGCGA ACAGCCAGGC AATTAATTGA GATTATTAAT 60

TGTACGAAAT ATATATGATG AGATTATAAA CACACAACAC TTTTATTCGC AAGGGATGAT 120

AAGAATCTAA TATATATATT ATCTGGGCTT CAAAGCATTG ATTTTATTTA TTGAGTCAAG 180

AGGGAAATTT ATTTTCTTGT TATTCTCTGT CCAGGTCTAA AGTCCCGAGC GGTGAGGCTA 240

TCTATTGATT TGGACATTCC AATCGAATAC AAAACAGAGA TACAGAAATT TGCGAAAAAA 300

TTTGATAACA ATCGTGGATT TTACGAATTA GACAAATTGA TATGTGCTTG CTAATTGATG 360

TGGCATGAAA TAAGAAATTT ATAAGGACGT TTTCAAGTGC TTCTATTTTA AACATTCAGG 420

ATTTTTTTTT AAAGCAGACA GCTTTCAACA GGTTTGATGA GAATTTGAAT ATTGATTGTT 480

GACTTTAGCT ATACATAAAT CACACCTCAT CCACCCATTG TGGTATCCTT CGAAGGACTT 540

GGGAACTGGA TCCTCTAGAA GGCGCGCC 568

(2) INFORMATION FOR SEQ ID NO: 3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 33 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

ATGCTAGCTC AGTACAATCC CAACACTCAG GCT 33

(2) INFORMATION FOR SEQ ID NO: 4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 32 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

CGAAGCTTAT AACTTGGCAT CAACGTGAAT TG 32

(2) INFORMATION FOR SEQ ID NO: 5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 56 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

CCATGGTTCT GGCCAAGAGC ATAGTGTGCC TCGCCCTCCT GGCGGTGGCG CTAGCT 56

(2) INFORMATION FOR SEQ ID NO: 6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

ATGCTAGCGC AATTCGACAC CAACTACGCA 30

(2) INFORMATION FOR SEQ ID NO: 7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

CAAGCTTGTT TACAACTTGG CGTTGACGT 29

(2) INFORMATION FOR SEQ ID NO: 8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1449 base pairs
(B) TYPE: nucleic acid

(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:

ATGCTAGCGC AATTCGACAC CAACTACGCA TCCGGTCGTA GTGGAATGGT CCACCTCTTC 60

GAGTGGAAGT GGGACGACAT CGCTGCCGAG TGCGAAAACT TCCTTGGACC CAATGGCTAC 120

GCCGGTGTTC AGGTCTCCCC TGTGAACGAG AACGCCGTCA AGGACAGCCG CCCCTGGTGG 180

GAACGTTACC AGCCCATCTC CTACAAGCTG GAGACCCGCT CCGGAAACGA AGAGCAGTTC 240

GCCAGCATGG TCAAGCGCTG CAACGCCGTC GGAGTGCGCA CCTACGTGGA CGTGGTCTTC 300

AACCACATGG CCGCCGACGG AGGCACCTAC GGCACTGGCG GCAGCACCGC CAGCCCCAGC 360

AGCAAGAGCT ATCCCGGAGT GCCCTACTCC TCGCTGGACT TCAACCCGAC CTGCGCCATC 420

AGCAACTACA ACGACGCCAA CGAGGTGCGC AACTGCGAGC TGGTCGGTCT GCGCGACCTT 480

AACCAGGGCA ACTCCTACGT GCAGGACAAG GTGGTCGAGT TCCTGGACCA TCTGATTGAT 540

CTCGGCGTGG CCGGATTCCG CGTGGACGCC GCCAAGCACA TGTGGCCCGC CGACCTGGCC 600

GTCATCTATG GCCGCCTCAA GAACCTAAAC ACCGACCACG GCTTCGCCTC GGGATCCAAG 660

GCGTACATCG TCCAGGAGGT CATCGACATG GGCGGCGAGG CCATCAGCAA GTCCGAGTAC 720

ACCGGACTGG GCGCCATCAC CGAGTTCCGC CACTCCGACT CCATCGGCAA GGTCTTCCGC 780

GGCAAGGACC AGCTGCAGTA CCTGACCAAC TGGGGCACCG CCTGGGGCTT CGCTGCCTCC 840

GACCGCTCCC TGGTATTCGT CGACAACCAC GACAATCAGC GCGGACATGG AGCAGGAGGC 900

GCCGACGTTC TGACCTACAA GGTGCCCAAG CAGTACAAGA TGGCCTCCGC CTTCATGCTG 960

GCGCACCCCT TCGGCACTCC CCGCGTGATG TCCTCCTTCT CCTTCACGGA CACCGATCAG 1020

GGCCCGCCCA CCACCGACGG CCACAACATC GCCTCGCCCA TCTTCAATAG CGACAACTCC 1080

TGCAGCGGCG GCTGGGTGTG TGAGCACCGC TGGCGCCAGA TCTACAACAT GGTGGCCTTC 1140

CGAAACACCG TGGGCTCGGA CGAGATCCAG AACTGGTGGG ACAACGGCAG CAACCAGATC 1200

TCCTTCAGCC GAGGCAGCCG CGGCTTCGTG GCCTTCAACA ACGACAACTA CGACCTGAAC 1260

AGCTCCCTGC AGACGGGCCT GCCCGCCGGC ACCTACTGCG ACGTCATCTC CGGCTCCAAG 1320

AGCGGTTCCT CCTGCACGGG CAAGACCGTC ACCGTCGGAT CCGACGGACG GGCTTCCATC 1380

AACATTGGCA GCTCCGAGGA CGACGGAGTG CTGGCCATTC ACGTCAACGC CAAGTTGTAA 1440

ACAAGCTTG 1449

(2) INFORMATION FOR SEQ ID NO: 9:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 7562 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:

GGTACCGAGC TCTTACGCGT GCTAGCCCGG GCTCGAGCGG CCGCCAGTGT GATGGATATC 60

TGCAGAATTC GGCTTAGTCG ACCCATGGTT CTGGCCAAGA GCATAGTGTG CCTCGCCCTC 120

CTGGCGGTGG CGCTAGCCCA ACGCCCAATT CGACACCAAC TACGCATCCG GTCGTAGTGG 180

AATGGTCCAC CTCTTCGAGT GGAAGTGGGA CGACATCGCT GCCGAGTGCG AAAACTTCCT 240

TGGACCCAAT GGCTACGCCG GTGTTCAGGT CTCCCCTGTG AACGAGAACG CCGTCAAGGA 300

CAGCCGCCCC TGGTGGGAAC GTTACCAGCC CATCTCCTAC AAGCTGGAGA CCCGCTCCGG 360

AAACGAAGAG CAGTTCGCCA GCATGGTCAA GCGCTGCAAC GCCGTCGGAG TGCGCACCTA 420

CGTGGACGTG GTCTTCAACC ACATGGCCGC CGACGGAGGC ACCTACGGCA CTGGCGGCAG 480

CACCGCCAGC CCCAGCAGCA AGAGCTATCC CGGAGTGCCC TACTCCTCGC TGGACTTCAA 540

CCCGACCTGC GCCATCAGCA ACTACAACGA CGCCAACGAG GTGCGCAACT GCGAGCTGGT 600

CGGTCTGCGC GACCTTAACC AGGGCAACTC CTACGTGCAG GACAAGGTGG TCGAGTTCCT 660

GGACCATCTG ATTGATCTCG GCGTGGCCGG ATTCCGCGTG GACGCCGCCA AGCACATGTG 720

GCCCGCCGAC CTGGCCGTCA TCTATGGCCG CCTCAAGAAC CTAAACACCG ACCACGGCTT 780

CGCCTCGGGA TCCAAGGCGT ACATCGTCCA GGAGGTCATC GACATGGGCG GCGAGGCCAT 840

CAGCAAGTCC GAGTACACCG GACTGGGCGC CATCACCGAG TTCCGCCACT CCGACTCCAT 900

CGGCAAGGTC TTCCGCGGCA AGGACCAGCT GCAGTACCTG ACCAACTGGG GCACCGCCTG 960

GGGCTTCGCT GCCTCCGACC GCTCCCTGGT ATTCGTCGAC AACCACGACA ATCAGCGCGG 1020

ACATGGAGCA GGAGGCGCCG ACGTTCTGAC CTACAAGGTG CCCAAGCAGT ACAAGATGGC 1080

CTCCGCCTTC ATGCTGGCGC ACCCCTTCGG CACTCCCCGC GTGATGTCCT CCTTCTCCTT 1140

CACGGACACC GATCAGGGCC CGCCCACCAC CGACGGCCAC AACATCGCCT CGCCCATCTT 1200

CAATAGCGAC AACTCCTGCA GCGGCGGCTG GGTGTGTGAG CACCGCTGGC GCCAGATCTA 1260

CAACATGGTG GCCTTCCGAA ACACCGTGGG CTCGGACGAG ATCCAGAACT GGTGGGACAA 1320

CGGCAGCAAC CAGATCTCCT TCAGCCGAGG CAGCCGCGGC TTCGTGGCCT TCAACAACGA 1380

CAACTACGAC CTGAACAGCT CCCTGCAGAC GGGCCTGCCC GCCGGCACCT ACTGCGACGT 1440

CATCTCCGGC TCCAAGAGCG GTTCCTCCTG CACGGGCAAG ACCGTCACCG TCGGATCCGA 1500

CGGACGGGCT TCCATCAACA TTGGCAGCTC CGAGGACGAC GGAGTGCTGG CCATTCACGT 1560

CAACGCCAAG TTGTAAGCTT GTAAACAGCT GGGGAGCATG GCGAACAGCC AGGCAATTAA 1620

TTGAGATTAT TAATTGTACG AAATATATAT GATGAGATTA TAAACACACA ACACTTTTAT 1680

TCGCAAGGGA TGATAAGAAT CTAATATATA TATTATCTGG GCTTCAAAGC ATTGATTTTA 1740

TTTATTGAGT CAAGAGGGAA ATTTATTTTC TTGTTATTCT CTGTCCAGGT CTAAAGTCCC 1800

GAGCGGTGAG GCTATCTATT GATTTGGACA TTCCAATCGA ATACAAAACA GAGATACAGA 1860

AATTTCGGAA AAAATTTGAT AACAATCGTG GATTTTAGCA ATTAGACAAA TTGATATGTG 1920

CTTGCTAATT GATGTGGCAT GAAATAAGAA ATTTATAAGG ACGTTTTCAA GTGCTTCTAT 1980

TTTAAACATT CAGGATTTTT TTTTAAAGCA GACAGCTTTC AACAGGTTTG ATCAGAATTT 2040

GAATATTGAT TGTTGACTTT AGCTATACAT AAATCACACC TCATCCACCC ATTGTGGTAT 2100

CCTTCGAAGG ACTTGGGAAC TGGATCCTCT AGAGCTTCCG GCCATTAGCC ATATTATTCA 2160

TTGGTTATAT AGCATAAATC AATATTGGCT ATTGGCCATT GCATACGTTG TATCCATATC 2220

ATAATATGTA CATTTATATT GGCTCATGTC CAACATTACC GCCATGTTGA CATTGATTAT 2280

TGACTAGTTA TTAATAGTAA TCAATTACGG GGTCATTAGT TCATAGCCCA TATATGGAGT 2340

TCCGCGTTAC ATAACTTACG GTAAATGGCC CGCCTGGCTG ACCGCCCAAC GACCCCCGCC 2400

CATTGACGTC AATAATGACG TATGTTCCCA TAGTAACGCC AATAGGGACT TTCCATTGAC 2460

GTCAATGGGT GGAGTATTTA CGGTAAACTG CCCACTTGGC AGTACATCAA GTGTATCATA 2520

TGCCAAGTAC GCCCCCTATT GACGTCAATG ACGGTAAATG GCCCGCCTGG CATTATGCCC 2580

AGTACATGAC CTTATGGGAC TTTCCTACTT GGCAGTACAT CTACGTATTA GTCATCGCTA 2640

TTACCATGGT GATGCGGTTT TGGCAGTACA TCAATGGGCG TGGATAGCGG TTTGACTCAC 2700

GGGGATTTCC AAGTCTCCAC CCCATTGACG TCAATGGGAG TTTGTTTTGG CACCAAAATC 2760

AACGGGACTT TCCAAAATGT CGTAACAACT CCGCCCCATT GACGCAAATG GGCGGTAGGC 2820

ATGTACGGTG GGAGGTCTAT ATAAGCAGAG CTCGTTTAGT GAACCGTCAG ATCGCCTGGA 2880

GACGCCATCC ACGCTGTTTT GACCTCCATA GAAGACACCG GGACCGATCC AGCCTCCGCG 2940

GCCCCAAGCT CATGGTTCTG GCCAAGAGCA TAGTGTGCCT CGCCCTCCTG GCGGTGGCGC 3000

TAGCCCAGTA CAATCCCAAC ACTCAGGCTG GGAGGACATC TATCGTGCAT CTCTTTGAAT 3060

GGCGCTGGGC CGACATTGCA CTGGAGTGCG AACACTATTT AGCTCCTAAT GGGTTTGGAG 3120

GAGTTCAGGT TTCTCCTCCA AATGAAAACA TTGTCATTAC TAATCCGAAC AGGCCCTGGT 3180

GGGAAAGATA CCAGCCCATC AGCTACAAGA TCTGCAGTCG ATCGGGCAAT GAAAATGAAT 3240

TCAGAGACAT GGTGACCAGA TGCAACAATG TTGGAGTTCG TATTTATGTG GATGCTGTTG 3300

TCAATCACAT GTGTGGATCT ATGGGTGGCA CGGGCACCCA CTCAACATGT GGGAGCTATT 3360

TCAACACCGG GACTAGAGAT TTTCCCGCTG TGCCGTACTC TGCCTGGGAT TTCAATGACG 3420

GCAAATGTCA CACTGCAAGT GGAGACATCG AAAATTATGG GGACATGTAT CAGGTCCGGG 3480

ATTGCAAGTT GTCCAGCCTT CTTGATCTGG CTCTGGAGAA GGACTATGTA CGCTCAACAA 3540

TTGCAGCGTA CATGAATCAC CTCATTGATA TGGGTGTAGC AGGGTTCCGG ATCGATGCTG 3600

CCAAGCATAT GTGGCCAGGG GACATAAGAG CGTTTCTGGA CAAACTGCAC GATCTAAATA 3660

CTCAGTGGTT TTCAGCAGGA ACGAAACCCT TTATTTACCA AGAGGTAATT GACTTGGGAG 3720

GAGAGCCAAT CACAGGCAGT CAGTACTTTG GGAATGGCCG CGTGACAGAA TTCAAGTATG 3780

GTGCCAAACT GGGGACGGTG ATCCGGAAGT GGAATGGAGA GAAGATGGCC TACTTAAAGA 3840

ACTGGGGAGA AGGCTGGGGC TTTGTGCCTT CTGACAGAGC CCTGGTGTTT GTGGATAACC 3900

ACGACAACCA GCGGGGGCAC GGGGCAGGCG GAGCTTCCAT TCTTACTTTC TGGGATGCCA 3960

GGCTTTATAA AATGGCGGTT GGTTTCATGC TCGCTCATCC GTACGGGTTC ACACGGGTGA 4020

TGTCAAGTTA TCGTTGGCCA AGATATTTCG AAAACGGAGT GGATGTTAAC GACTGGGTGG 4080

GACCACCAAG TAACTCGGAC GGATCGACGA AGTCCGTTAC AATCAACGCA GACACTACCT 4140

GTGGCAATGA CTGGGTCTGC GAACATCGCT GGCGACAAAT AAGGAACATG GTTATCTTCC 4200

GTAATGTGGT AGACGGTCAG CCTTTCTCAA ACTGGTGGGA CAACGGGAGC AATCAAGTAG 4260

CTTTCGGTCG CGGCGACAGA GGCTTCATTG TCTTTAATAA TGATGACTGG TATATGAATG 4320

TCGATTTGCA AACTGGTCTG CCTGCTGGAA CCTACTGCGA TGTTATTTCT GGACAAAAGG 4380

AAGGCAGTGC GTGTACTGGA AAGCAGGTGT ACGTTTCCTC GGATGGAAAG GCCAATTTCC 4440

AGATTAGTAA CAGCGATGAA GATCCATTTG TTGCAATTCA CGTTGATGCC AAGTTATAAG 4500

GCCGGCCGCT TCGAGCAGAC ATGATAAGAT ACATTGATGA GTTTGGACAA ACCACAACTA 4560

GAATGCAGTG AAAAAAATGC TTTATTTGTG AAATTTGTGA TGCTATTGCT TTATTTGTAA 4620

CCATTATAAG CTGCAATAAA CAAGTTAACA ACAACAATTG CATTCATTTT ATGTTTCAGG 4680

TTCAGGGGGA GGTGTGGGAG GTTTTTTAAA GCAAGTAAAA CCTCTACAAA TGTGGTAAAA 4740

TCGATAAGGA TCCGTCGACC GATGCCCTTG AGAGCCTTCA ACCCAGTCAG CTCCTTCCGG 4800

TGGGCGCGGG GCATGACTAT CGTCGCCGCA CTTATGACTG TCTTCTTTAT CATGCAACTC 4860

GTAGGACAGG TGCCGGCAGC GCTCTTCCGC TTCCTCGCTC ACTGACTCGC TGCGCTCGGT 4920

CGTTCGGCTG CGGCGAGCGG TATCAGCTCA CTCAAAGGCG GTAATACGGT TATCCACAGA 4980

ATCAGGGGAT AACGCAGGAA AGAACATGTG AGCAAAAGGC CAGCAAAAGG CCAGGAACCG 5040

TAAAAAGGCC GCGTTGCTGG CGTTTTTCCA TAGGCTCCGC CCCCCTGACG AGCATCACAA 5100

AAATCGACGC TCAAGTCAGA GGTGGCGAAA CCCGACAGGA CTATAAAGAT ACCAGGCGTT 5160

TCCCCCTGGA AGCTCCCTCG TGCGCTCTCC TGTTCCGACC CTGCCGCTTA CCGGATACCT 5220

GTCCGCCTTT CTCCCTTCGG GAAGCGTGGC GCTTTCTCAT AGCTCACGCT GTAGGTATCT 5280

CAGTTCGGTG TAGGTCGTTC GCTCCAAGCT GGGCTGTGTG CACGAACCCC CCGTTCAGCC 5340

CGACCGCTGC GCCTTATCCG GTAACTATCG TCTTGAGTCC AACCCGGTAA GACACGACTT 5400

ATCGCCACTG GCAGCAGCCA CTGGTAACAG GATTAGCAGA GCGAGGTATG TAGGCGGTGC 5460

TACAGAGTTC TTGAAGTGGT GGCCTAACTA CGGCTACACT AGAAGGACAG TATTTGGTAT 5520

CTGCGCTCTG CTGAAGCCAG TTACCTTCGG AAAAAGAGTT GGTAGCTCTT GATCCGGCAA 5580

ACAAACCACC GCTGGTAGCG GTGGTTTTTT TGTTTGCAAG CAGCAGATTA CGCGCAGAAA 5640

AAAAGGATCT CAAGAAGATC CTTTGATCTT TTCTACGGGG TCTGACGCTC AGTGGAACGA 5700

AAACTCACGT TAAGGGATTT TGGTCATGAG ATTATCAAAA AGGATCTTCA CCTAGATCCT 5760

TTTAAATTAA AAATGAAGTT TTAAATCAAT CTAAAGTATA TATGAGTAAA CTTGGTCTGA 5820

CAGTTACCAA TGCTTAATCA GTGAGGCACC TATCTCAGCG ATCTGTCTAT TTCGTTCATC 5880

CATAGTTGCC TGACTCCCCG TCGTGTAGAT AACTACGATA CGGGAGGGCT TACCATCTGG 5940

CCCCAGTGCT GCAATGATAC CGCGAGACCC ACGCTCACCG GCTCCAGATT TATCAGCAAT 6000

AAACCAGCCA GCCGGAAGGG CCGAGCGCAG AAGTGGTCCT GCAACTTTAT CCGCCTCCAT 6060

CCAGTCTATT AATTGTTGCC GGGAAGCTAG AGTAAGTAGT TCGCCAGTTA ATAGTTTGCG 6120

CAACGTTGTT GCCATTGCTA CAGGCATCGT GGTGTCACGC TCGTCGTTTG GTATGGCTTC 6180

ATTCAGCTCC GGTTCCCAAC GATCAAGGCG AGTTACATGA TCCCCCATGT TGTGCAAAAA 6240

AGCGGTTAGC TCCTTCGGTC CTCCGATCGT TGTCAGAAGT AAGTTGGCCG CAGTGTTATC 6300

ACTCATGGTT ATGGCAGCAC TGCATAATTC TCTTACTGTC ATGCCATCCG TAAGATGCTT 6360

TTCTGTGACT GGTGAGTACT CAACCAAGTC ATTCTGAGAA TAGTGTATGC GGCGACCGAG 6420

TTGCTCTTGC CCGGCGTCAA TACGGGATAA TACCGCGCCA CATAGCAGAA CTTTAAAAGT 6480

GCTCATCATT GGAAAACGTT CTTCGGGGCG AAAACTCTCA AGGATCTTAC CGCTGTTGAG 6540

ATCCAGTTCG ATGTAACCCA CTCGTGCACC CAACTGATCT TCAGCATCTT TTACTTTCAC 6600

CAGCGTTTCT GGGTGAGCAA AAACAGGAAG GCAAAATGCC GCAAAAAAGG GAATAAGGGC 6660

GACACGGAAA TGTTGAATAC TCATACTCTT CCTTTTTCAA TATTATTGAA GCATTTATCA 6720

GGGTTATTGT CTCATGAGCG GATACATATT TGAATGTATT TAGAAAAATA AACAAATAGG 6780

GGTTCCGCGC ACATTTCCCC GAAAAGTGCC ACCTGACGCG CCCTGTAGCG GCGCATTAAG 6840

CGCGGCGGGT GTGGTGGTTA CGCGCAGCGT GACCGCTACA CTTGCCAGCG CCCTAGCGCC 6900

CGCTCCTTTC GCTTTCTTCC CTTCCTTTCT CGCCACGTTC GCCGGCTTTC CCCGTCAAGC 6960

TCTAAATCGG GGGCTCCCTT TAGGGTTCCG ATTTAGTGCT TTACGGCACC TCGACCCCAA 7020

AAAACTTGAT TAGGGTGATG GTTCACGTAG TGGGCCATCG CCCTGATAGA CGGTTTTTCG 7080

CCCTTTGACG TTGGAGTCCA CGTTCTTTAA TAGTGGACTC TTGTTCCAAA CTGGAACAAC 7140

ACTCAACCCT ATCTCGGTCT ATTCTTTTGA TTTATAAGGG ATTTTGCCGA TTTCGGCCTA 7200

TTGGTTAAAA AATGAGCTGA TTTAACAAAA ATTTAACGCG AATTTTAACA AAATATTAAC 7260

GTTTACAATT TCCCATTCGC CATTCAGGCT GCGCAACTGT TGGGAAGGGC GATCGGTGCG 7320

GGCCTCTTCG CTATTACGCC AGCCCAAGCT ACCATGATAA GTAAGTAATA TTAAGGTACG 7380

GGAGGTACTT GGAGCGGCCG CAATAAAATA TCTTTATTTT CATTACATCT GTGTGTTGGT 7440

TTTTTGTGTG AATCGATAGT ACTAACATAC GCTCTCCATC AAAACAAAAC GAAACAAAAC 7500

AAACTAGCAA AATAGGCTGT CCCCAGTGCA AGTGCAGGTG CCAGAACATT TCTCTATCGA 7560

TA 7562

(2) INFORMATION FOR SEQ ID NO: 10:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 7913 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:

GGTACCGAAA CGCGCGAGGC AGCTGATCTG AGTCACTGAA TGGGATAGTG TCCACAAGGG 60

CGGGGGCTAT TCTTGTCCAT CTGGGCCTAC GGAACCAGCA CCCATCGCCA AACTCTTCAT 120

CCTCTTCCTC AATCTCGCTT TCTCTCTCGT CCGCTTTTTT TTTCTTCTTC TTTTTTTTTT 180

TTTTTTTTTT TTTTTTTTGC AAAAGGAGGG GAGAGGGGGT AAAAAAATGC TGCACTGTGC 240

GGCGAGGCCG GTGAGTGAGC GACGCGGAGC CAATCAGCGC CCGCCGTTCC GAAAGTTGCC 300

TTTTATGGCT CGAGTGGCCG CTGTGGCGTC CTATAAAACC CGGCGGCGCA ACGCGCAGCC 360

ACTGTCGAGT CAAGCTTCGA TCGTCGAGCG GCCGCCAGTG TGATGGATAT CTGCAGAATT 420

CGGCTTAGTC GACCCATGGT TCTGGCCAAG AGCATAGTGT GCCTCGCCCT CCTGGCGGTG 480

GCGCTAGCCC AACGCCCAAT TCGACACCAA CTACGCATCC GGTCGTAGTG GAATGGTCCA 540

CCTCTTCGAG TGGAAGTGGG ACGACATCGC TGCCGAGTGC GAAAACTTCC TTGGACCCAA 600

TGGCTACGCC GGTGTTCAGG TCTCCCCTGT GAACGAGAAC GCCGTCAAGG ACAGCCGCCC 660

CTGGTGGGAA CGTTACCAGC CCATCTCCTA CAAGCTGGAG ACCCGCTCCG GAAACGAAGA 720

GCAGTTCGCC AGCATGGTCA AGCGCTGCAA CGCCGTCGGA GTGCGCACCT ACGTGGACGT 780

GGTCTTCAAC CACATGGCCG CCGACGGAGG CACCTACGGC ACTGGCGGCA GCACCGCCAG 840

CCCCAGCAGC AAGAGCTATC CCGGAGTGCC CTACTCCTCG CTGGACTTCA ACCCGACCTG 900

CGCCATCAGC AACTACAACG ACGCCAACGA GGTGCGCAAC TGCGAGCTGG TCGGTCTGCG 960

CGACCTTAAC CAGGGCAACT CCTACGTGCA GGACAAGGTG GTCGAGTTCC TGGACCATCT 1020

GATTGATCTC GGCGTGGCCG GATTCCGCGT GGACGCCGCC AAGCACATGT GGCCCGCCGA 1080

CCTGGCCGTC ATCTATGGCC GCCTCAAGAA CCTAAACACC GACCACGGCT TCGCCTCGGG 1140

ATCCAAGGCG TACATCGTCC AGGAGGTCAT CGACATGGGC GGCGAGGCCA TCAGCAAGTC 1200

CGAGTACACC GGACTGGGCG CCATCACCGA GTTCCGCCAC TCCGACTCCA TCGGCAAGGT 1260

CTTCCGCGGC AAGGACCAGC TGCAGTACCT GACCAACTGG GGCACCGCCT GGGGCTTCGC 1320

TGCCTCCGAC CGCTCCCTGG TATTCGTCGA CAACCACGAC AATCAGCGCG GACATGGAGC 1380

AGGAGGCGCC GACGTTCTGA CCTACAAGGT GCCCAAGCAG TACAAGATGG CCTCCGCCTT 1440

CATGCTGGCG CACCCCTTCG GCACTCCCCG CGTGATGTCC TCCTTCTCCT TCACGGACAC 1500

CGATCAGGGC CCGCCCACCA CCGACGGCCA CAACATCGCC TCGCCCATCT TCAATAGCGA 1560

CAACTCCTGC AGCGGCGGCT GGGTGTGTGA GCACCGCTGG CGCCAGATCT ACAACATGGT 1620

GGCCTTCCGA AACACCGTGG GCTCGGACGA GATCCAGAAC TGGTGGGACA ACGGCAGCAA 1680

CCAGATCTCC TTCAGCCGAG GCAGCCGCGG CTTCGTGGCC TTCAACAACG ACAACTACGA 1740

CCTGAACAGC TCCCTGCAGA CGGGCCTGCC CGCCGGCACC TACTGCGACG TCATCTCCGG 1800

CTCCAAGAGC GGTTCCTCCT GCACGGGCAA GACCGTCACC GTCGGATCCG ACGGACGGGC 1860

TTCCATCAAC ATTGGCAGCT CCGAGGACGA CGGAGTGCTG GCCATTCACG TCAACGCCAA 1920

GTTGTAAGCT TGTAAACAGC TGGGGAGCAT GGCGAACAGC CAGGCAATTA ATTGAGATTA 1980

TTAATTGTAC GAAATATATA TGATGAGATT ATAAACACAC AACACTTTTA TTCGCAAGGG 2040

ATGATAAGAA TCTAATATAT ATATTATCTG GGCTTCAAAG CATTGATTTT ATTTATTGAG 2100

TCAAGAGGGA AATTTATTTT CTTGTTATTC TCTGTCCAGG TCTAAAGTCC CGAGCGGTGA 2160

GGCTATCTAT TGATTTGGAC ATTCCAATCG AATACAAAAC AGAGATACAG AAATTTCGGA 2220

AAAAATTTGA TAACAATCGT GGATTTTAGC AATTAGACAA ATTGATATGT GCTTGCTAAT 2280

TGATGTGGCA TGAAATAAGA AATTTATAAG GACGTTTTCA AGTGCTTCTA TTTTAAACAT 2340

TCAGGATTTT TTTTTAAAGC AGACAGCTTT CAACAGGTTT GATCAGAATT TGAATATTGA 2400

TTGTTGACTT TAGCTATACA TAAATCACAC CTCATCCACC CATTGTGGTA TCCTTCGAAG 2460

GACTTGGGAA CTGGATCCTC TAGAGCTTCC GGCCATTAGC CATATTATTC ATTGGTTATA 2520

TAGCATAAAT CAATATTGGC TATTGGCCAT TGCATACGTT GTATCCATAT CATAATATGT 2580

ACATTTATAT TGGCTCATGT CCAACATTAC CGCCATGTTG ACATTGATTA TTGACTAGTT 2640

ATTAATAGTA ATCAATTACG GGGTCATTAG TTCATAGCCC ATATATGGAG TTCCGCGTTA 2700

CATAACTTAC GGTAAATGGC CCGCCTGGCT GACCGCCCAA CGACCCCCGC CCATTGACGT 2760

CAATAATGAC GTATGTTCCC ATAGTAACGC CAATAGGGAC TTTCCATTGA CGTCAATGGG 2820

TGGAGTATTT ACGGTAAACT GCCCACTTGG CAGTACATCA AGTGTATCAT ATGCCAAGTA 2880

CGCCCCCTAT TGACGTCAAT GACGGTAAAT GGCCCGCCTG GCATTATGCC CAGTACATGA 2940

CCTTATGGGA CTTTCCTACT TGGCAGTACA TCTACGTATT AGTCATCGCT ATTACCATGG 3000

TGATGCGGTT TTGGCAGTAC ATCAATGGGC GTGGATAGCG GTTTGACTCA CGGGGATTTC 3060

CAAGTCTCCA CCCCATTGAC GTCAATGGGA GTTTGTTTTG GCACCAAAAT CAACGGGACT 3120

TTCCAAAATG TCGTAACAAC TCCGCCCCAT TGACGCAAAT GGGCGGTAGG CATGTACGGT 3180

GGGAGGTCTA TATAAGCAGA GCTCGTTTAG TGAACCGTCA GATCGCCTGG AGACGCCATC 3240

CACGCTGTTT TGACCTCCAT AGAAGACACC GGGACCGATC CAGCCTCCGC GGCCCCAAGC 3300

TCATGGTTCT GGCCAAGAGC ATAGTGTGCC TCGCCCTCCT GGCGGTGGCG CTAGCCCAGT 3360

ACAATCCCAA CACTCAGGCT GGGAGGACAT CTATCGTGCA TCTCTTTGAA TGGCGCTGGG 3420

CCGACATTGC ACTGGAGTGC GAACACTATT TAGCTCCTAA TGGGTTTGGA GGAGTTCAGG 3480

TTTCTCCTCC AAATGAAAAC ATTGTCATTA CTAATCCGAA CAGGCCCTGG TGGGAAAGAT 3540

ACCAGCCCAT CAGCTACAAG ATCTGCAGTC GATCGGGCAA TGAAAATGAA TTCAGAGACA 3600

TGGTGACCAG ATGCAACAAT GTTGGAGTTC GTATTTATGT GGATGCTGTT GTCAATCACA 3660

TGTGTGGATC TATGGGTGGC ACGGGCACCC ACTCAACATG TGGGAGCTAT TTCAACACCG 3720

GGACTAGAGA TTTTCCCGCT GTGCCGTACT CTGCCTGGGA TTTCAATGAC GGCAAATGTC 3780

ACACTGCAAG TGGAGACATC GAAAATTATG GGGACATGTA TCAGGTCCGG GATTGCAAGT 3840

TGTCCAGCCT TCTTGATCTG GCTCTGGAGA AGGACTATGT ACGCTCAACA ATTGCAGCGT 3900

ACATGAATCA CCTCATTGAT ATGGGTGTAG CAGGGTTCCG GATCGATGCT GCCAAGCATA 3960

TGTGGCCAGG GGACATAAGA GCGTTTCTGG ACAAACTGCA CGATCTAAAT ACTCAGTGGT 4020

TTTCAGCAGG AACGAAACCC TTTATTTACC AAGAGGTAAT TGACTTGGGA GGAGAGCCAA 4080

TCACAGGCAG TCAGTACTTT GGGAATGGCC GCGTGACAGA ATTCAAGTAT GGTGCCAAAC 4140

TGGGGACGGT GATCCGGAAG TGGAATGGAG AGAAGATGGC CTACTTAAAG AACTGGGGAG 4200

AAGGCTGGGG CTTTGTGCCT TCTGACAGAG CCCTGGTGTT TGTGGATAAC CACGACAACC 4260

AGCGGGGGCA CGGGGCAGGC GGAGCTTCCA TTCTTACTTT CTGGGATGCC AGGCTTTATA 4320

AAATGGCGGT TGGTTTCATG CTCGCTCATC CGTACGGGTT CACACGGGTG ATGTCAAGTT 4380

ATCGTTGGCC AAGATATTTC GAAAACGGAG TGGATGTTAA CGACTGGGTG GGACCACCAA 4440

GTAACTCGGA CGGATCGACG AAGTCCGTTA CAATCAACGC AGACACTACC TGTGGCAATG 4500

ACTGGGTCTG CGAACATCGC TGGCGACAAA TAAGGAACAT GGTTATCTTC CGTAATGTGG 4560

TAGACGGTCA GCCTTTCTCA AACTGGTGGG ACAACGGGAG CAATCAAGTA GCTTTCGGTC 4620

GCGGCGACAG AGGCTTCATT GTCTTTAATA ATGATGACTG GTATATGAAT GTCGATTTGC 4680

AAACTGGTCT GCCTGCTGGA ACCTACTGCG ATGTTATTTC TGGACAAAAG GAAGGCAGTG 4740

CGTGTACTGG AAAGCAGGTG TACGTTTCCT CGGATGGAAA GGCCAATTTC CAGATTAGTA 4800

ACAGCGATGA AGATCCATTT GTTGCAATTC ACGTTGATGC CAAGTTATAA GGCCGGCCGC 4860

TTCGAGCAGA CATGATAAGA TACATTGATG AGTTTGGACA AACCACAACT AGAATGCAGT 4920

GAAAAAAATG CTTTATTTGT GAAATTTGTG ATGCTATTGC TTTATTTGTA ACCATTATAA 4980

GCTGCAATAA ACAAGTTAAC AACAACAATT GCATTCATTT TATGTTTCAG GTTCAGGGGG 5040

AGGTGTGGGA GGTTTTTTAA AGCAAGTAAA ACCTCTACAA ATGTGGTAAA ATCGATAAGG 5100

ATCCGTCGAC CGATGCCCTT GAGAGCCTTC AACCCAGTCA GCTCCTTCCG GTGGGCGCGG 5160

GGCATGACTA TCGTCGCCGC ACTTATGACT GTCTTCTTTA TCATGCAACT CGTAGGACAG 5220

GTGCCGGCAG CGCTCTTCCG CTTCCTCGCT CACTGACTCG CTGCGCTCGG TCGTTCGGCT 5280

GCGGCGAGCG GTATCAGCTC ACTCAAAGGC GGTAATACGG TTATCCACAG AATCAGGGGA 5340

TAACGCAGGA AAGAACATGT GAGCAAAAGG CCAGCAAAAG GCCAGGAACC GTAAAAAGGC 5400

CGCGTTGCTG GCGTTTTTCC ATAGGCTCCG CCCCCCTGAC GAGCATCACA AAAATCGACG 5460

CTCAAGTCAG AGGTGGCGAA ACCCGACAGG ACTATAAAGA TACCAGGCGT TTCCCCCTGG 5520

AAGCTCCCTC GTGCGCTCTC CTGTTCCGAC CCTGCCGCTT ACCGGATACC TGTCCGCCTT 5580

TCTCCCTTCG GGAAGCGTGG CGCTTTCTCA TAGCTCACGC TGTAGGTATC TCAGTTCGGT 5640

GTAGGTCGTT CGCTCCAAGC TGGGCTGTGT GCACGAACCC CCCGTTCAGC CCGACCGCTG 5700

CGCCTTATCC GGTAACTATC GTCTTGAGTC CAACCCGGTA AGACACGACT TATCGCCACT 5760

GGCAGCAGCC ACTGGTAACA GGATTAGCAG AGCGAGGTAT GTAGGCGGTG CTACAGAGTT 5820

CTTGAAGTGG TGGCCTAACT ACGGCTACAC TAGAAGGACA GTATTTGGTA TCTGCGCTCT 5880

GCTGAAGCCA GTTACCTTCG GAAAAAGAGT TGGTAGCTCT TGATCCGGCA AACAAACCAC 5940

CGCTGGTAGC GGTGGTTTTT TTGTTTGCAA GCAGCAGATT ACGCGCAGAA AAAAAGGATC 6000

TCAAGAAGAT CCTTTGATCT TTTCTACGGG GTCTGACGCT CAGTGGAACG AAAACTCACG 6060

TTAAGGGATT TTGGTCATGA GATTATCAAA AAGGATCTTC ACCTAGATCC TTTTAAATTA 6120

AAAATGAAGT TTTAAATCAA TCTAAAGTAT ATATGAGTAA ACTTGGTCTG ACAGTTACCA 6180

ATGCTTAATC AGTGAGGCAC CTATCTCAGC GATCTGTCTA TTTCGTTCAT CCATAGTTGC 6240

CTGACTCCCC GTCGTGTAGA TAACTACGAT ACGGGAGGGC TTACCATCTG GCCCCAGTGC 6300

TGCAATGATA CCGCGAGACC CACGCTCACC GGCTCCAGAT TTATCAGCAA TAAACCAGCC 6360

AGCCGGAAGG GCCGAGCGCA GAAGTGGTCC TGCAACTTTA TCCGCCTCCA TCCAGTCTAT 6420

TAATTGTTGC CGGGAAGCTA GAGTAAGTAG TTCGCCAGTT AATAGTTTGC GCAACGTTGT 6480

TGCCATTGCT ACAGGCATCG TGGTGTCACG CTCGTCGTTT GGTATGGCTT CATTCAGCTC 6540

CGGTTCCCAA CGATCAAGGC GAGTTACATG ATCCCCCATG TTGTGCAAAA AAGCGGTTAG 6600

CTCCTTCGGT CCTCCGATCG TTGTCAGAAG TAAGTTGGCC GCAGTGTTAT CACTCATGGT 6660

TATGGCAGCA CTGCATAATT CTCTTACTGT CATGCCATCC GTAAGATGCT TTTCTGTGAC 6720

TGGTGAGTAC TCAACCAAGT CATTCTGAGA ATAGTGTATG CGGCGACCGA GTTGCTCTTG 6780

CCCGGCGTCA ATACGGGATA ATACCGCGCC ACATAGCAGA ACTTTAAAAG TGCTCATCAT 6840

TGGAAAACGT TCTTCGGGGC GAAAACTCTC AAGGATCTTA CCGCTGTTGA GATCCAGTTC 6900

GATGTAACCC ACTCGTGCAC CCAACTGATC TTCAGCATCT TTTACTTTCA CCAGCGTTTC 6960

TGGGTGAGCA AAAACAGGAA GGCAAAATGC CGCAAAAAAG GGAATAAGGG CGACACGGAA 7020

ATGTTGAATA CTCATACTCT TCCTTTTTCA ATATTATTGA AGCATTTATC AGGGTTATTG 7080

TCTCATGAGC GGATACATAT TTGAATGTAT TTAGAAAAAT AAACAAATAG GGGTTCCGCG 7140

CACATTTCCC CGAAAAGTGC CACCTGACGC GCCCTGTAGC GGCGCATTAA GCGCGGCGGG 7200

TGTGGTGGTT ACGCGCAGCG TGACCGCTAC ACTTGCCAGC GCCCTAGCGC CCGCTCCTTT 7260

CGCTTTCTTC CCTTCCTTTC TCGCCACGTT CGCCGGCTTT CCCCGTCAAG CTCTAAATCG 7320

GGGGCTCCCT TTAGGGTTCC GATTTAGTGC TTTACGGCAC CTCGACCCCA AAAAACTTGA 7380

TTAGGGTGAT GGTTCACGTA GTGGGCCATC GCCCTGATAG ACGGTTTTTC GCCCTTTGAC 7440

GTTGGAGTCC ACGTTCTTTA ATAGTGGACT CTTGTTCCAA ACTGGAACAA CACTCAACCC 7500

TATCTCGGTC TATTCTTTTG ATTTATAAGG GATTTTGCCG ATTTCGGCCT ATTGGTTAAA 7560

AAATGAGCTG ATTTAACAAA AATTTAACGC GAATTTTAAC AAAATATTAA CGTTTACAAT 7620

TTCCCATTCG CCATTCAGGC TGCGCAACTG TTGGGAAGGG CGATCGGTGC GGGCCTCTTC 7680

GCTATTACGC CAGCCCAAGC TACCATGATA AGTAAGTAAT ATTAAGGTAC GGGAGGTACT 7740

TGGAGCGGCC GCAATAAAAT ATCTTTATTT TCATTACATC TGTGTGTTGG TTTTTTGTGT 7800

GAATCGATAG TACTAACATA CGCTCTCCAT CAAAACAAAA CGAAACAAAA CAAACTAGCA 7860

AAATAGGCTG TCCCCAGTGC AAGTGCAGGT GCCAGAACAT TTCTCTATCG ATA 7913

THE EMBODIMENTS OF THE INVENTION IN WHICH AN EXCLUSIVE PROPERTY OF PRIVILEGE IS CLAIMED ARE DEFINED AS FOLLOWS:

1. An extra-cellular reporter gene system for use in a target cell, comprising a DNA sequence encoding:

a signal peptide;

at least one α-amylase protein; and

a transcription termination region.

2. The reporter gene system of claim 1, wherein the α-amylase protein is electrophoretically distinct from any endogenous amylase in the target cell and wherein the α-amylase protein is compatible with the target cells expression and secretion systems.

3. The reporter gene system of claim 2, wherein the reporter gene system comprises DNA sequences encoding two α-amylase proteins.

4. The reporter gene system of claim 3, wherein the α-amylase proteins are electrophoretically distinct from each other.

5. The reporter gene system of claim 2, wherein the α-amylase protein is selected from the group consisting of chicken α-amylase and *Drosophila melanogaster* α-amylase.

6. The reporter gene system of Claim 5, wherein the signal peptide is derived from a *D. melanogaster* α-amylase gene.

7. The reporter gene system of Claim 6, wherein the signal peptide is encoded by the DNA sequence of SEQID No: 5.

8. The reporter gene system of Claim 7, wherein the chicken α-amylase protein is encoded by the DNA sequence of SEQID No: 1.

9. The reporter gene system of Claim 7, wherein the *D. melanogaster* α-amylase protein is encoded by the DNA sequence of SEQID No: 8.

10. The reporter gene system of Claim 7, wherein the transcription termination region is derived from the group consisting of a *D. melanogaster* α-amylase gene and a late polyadenylation site of SV40.

11. The reporter gene system of Claim 10, wherein the transcription termination region is derived from the *D. melanogaster* α-amylase gene and is encoded by the DNA sequence of SEQID No: 2.

12. The reporter gene system of Claim 1, wherein said system further comprises a promoter region.

13. The reporter gene system of Claim 12, wherein the promoter region is selected from the group consisting of: a Cytomegalovirus promoter, a murine retroviral promoter and a rat β-actin promoter.

14. The reporter gene system of claim 12, wherein the α-amylase protein is electrophoretically distinct from any endogenous amylase in the target cell and wherein the α-amylase protein is compatible with the target cells expression and secretion systems.

15. The reporter gene system of claim 14, wherein the reporter gene system comprises DNA sequences encoding two α-amylase proteins.

16. The reporter gene system of claim 15, wherein the α-amylase proteins are electrophoretically distinct from each other.

17. The reporter gene system of claim 14, wherein the α-amylase protein is selected from the group consisting of chicken α-amylase and *Drosophila melanogaster* α-amylase.

18. The reporter gene system of Claim 17, wherein the signal peptide is derived from a *D. melanogaster* α-amylase gene.

19. The reporter gene system of Claim 18, wherein the signal peptide is encoded by the DNA sequence of SEQID No: 5.

20. The reporter gene system of Claim 17, wherein the chicken α-amylase protein is encoded by the DNA sequence of SEQID No: 1.

21. The reporter gene system of Claim 17, wherein the *D. melanogaster* α-amylase protein is encoded by the DNA sequence of SEQID No: 8.

22. The reporter gene system of Claim 17, wherein the transcription termination region is derived from the group consisting of a *D. melanogaster* α-amylase gene and a late polyadenylation site of SV40.

23. The reporter gene system of Claim 22, wherein the transcription termination region is derived from the *D. melanogaster* α-amylase gene and is encoded by the DNA sequence of SEQID No: 2.

24. The reporter gene system of Claim 11, wherein the extra-cellular reporter gene system comprising a DNA sequence encoding:

a signal peptide derived from a *Drosophila melanogaster* α-amylase gene;

a chicken α-amylase protein; and

a transcription termination region derived from a *Drosophila melanogaster* α-amylase gene.

25. The reporter gene system of Claim 23, wherein the extra-cellular reporter gene system comprising a DNA sequence encoding:

a Cytomegalovirus promoter,

a signal peptide derived from a *Drosophila melanogaster* α-amylase gene;

a chicken α-amylase protein; and

a transcription termination region derived from a *Drosophila melanogaster* α-amylase gene.

26. The reporter gene system of Claim 11, wherein the reporter gene system comprising a DNA sequence encoding:

a signal peptide derived from a *Drosophila melanogaster* α-amylase gene;

a Drosophila α-amylase protein; and

a transcription termination region derived from a *Drosophila melanogaster* α-amylase gene.

27. The reporter gene system of Claim 23, wherein the extra-cellular reporter gene system comprising a DNA sequence encoding:

a Cytomegalovirus promoter,

a signal peptide derived from a *Drosophila melanogaster* α-amylase gene;

a *Drosophila* α-amylase protein; and

a transcription termination region derived from a *Drosophila melanogaster* α-amylase gene.

28. The reporter gene system of Claim 16, wherein the extra-cellular reporter gene system comprising a DNA sequence encoding:

two expression cassettes, wherein the first cassette comprises a DNA sequence encoding:

a signal peptide derived from a *Drosophila melanogaster* α-amylase gene;

a *Drosophila* α-amylase protein; and

a transcription termination region derived from a *Drosophila melanogaster* α-amylase gene;

and wherein said second cassette comprises a DNA sequence encoding:

a Cytomegalovirus promoter,

a signal peptide derived from a *D. melanogaster* α-amylase gene;

a chicken α-amylase protein; and

a transcription termination region derived from a late polyadenylation site of SV40;

wherein the first cassette further comprises a polylinker region immediately 3' of the *D. melanogaster* α-amylase gene to facilitate insertion of a promoter sequence of interest to determine the strength of said promoter.

29. The reporter gene system of any of the preceding claims were the target cell is a mammalian cell.

30. A kit comprising the reporter gene system of any of the preceding claims, a means of detecting the amylase activity and instructions for use.

31. The kit of claim 30 further comprising a beta-dextrin system to remove non-specific proteins from high-protein fluids.

32. The kit of claim 30 or 31, wherein the means of detecting the amylase system is selected from the group consisting of a plate assay and a gel electrophoresis means.

5'-GCTAGCTCAG TACAATCCCA ACACTCAGGC TGGGAGGACA TCTATCGTGC ATCTCTTTGA
ATGGCGCTGG GCCGACATTG CACTGGAGTG CGAACACTAT TTAGCTCCTA ATGGGTTTGG
AGGAGTTCAG GTTTCTCCTC CAAATGAAAA CATTGTCATT ACTAATCCGA ACAGGCCCTG
GTGGGAAAGA TACCAGCCCA TCAGCTACAA GATCTGCAGT CGATCGGGCA ATGAAAATGA
ATTCAGAGAC ATGGTGACCA GATGCAACAA TGTTGGAGTT CGTATTTATG TGGATGCTGT
TGTCAATCAC ATGTGTGGAT CTATGGGTGG CACGGGCACC CACTCAACAT GTGGGAGCTA
TTTCAACACC GGGACTAGAG ATTTTCCCGC TGTGCCGTAC TCTGCCTGGG ATTTCAATGA
CGGCAAATGT CACACTGCAA GTGGAGACAT CGAAAATTAT GGGGACATGT ATCAGGTCCG
GGATTGCAAG TTGTCCAGCC TTCTTGATCT GGCTCTGGAG AAGGACTATG TACGCTCAAC
AATTGCAGCG TACATGAATC ACCTCATTGA TATGGGTGTA GCAGGGTTCC GGATCGATGC
TGCCAAGCAT ATGTGGCCAG GGGACATAAG AGCGTTTCTG GACAAACTGC ACGATCTAAA
TACTCAGTGG TTTTCAGCAG GAACGAAACC CTTTATTTAC CAAGAGGTAA TTGACTTGGG
AGGAGAGCCA ATCACAGGCA GTCAGTACTT TGGGAATGGC CGCGTGACAG AATTCAAGTA
TGGTGCCAAA CTGGGGACGG TGATCCGGAA GTGGAATGGA GAGAAGATGG CCTACTTAAA
GAACTGGGGA GAAGGCTGGG GCTTTGTGCC TTCTGACAGA GCCCTGGTGT TTGTGGATAA
CCACGACAAC CAGCGGGGGC ACGGGGCAGG CGGAGCTTCC ATTCTTACTT TCTGGGATGC
CAGGCTTTAT AAAATGGCGG TTGGTTTCAT GCTCGCTCAT CCGTACGGGT TCACACGGGT
GATGTCAAGT TATCGTTGGC CAAGATATTT CGAAAACGGA GTGGATGTTA ACGACTGGGT
GGGACCACCA AGTAACTCGG ACGGATCGAC GAAGTCCGTT ACAATCAACG CAGACACTAC
CTGTGGCAAT GACTGGGTCT GCGAACATCG CTGGCGACAA ATAAGGAACA TGGTTATCTT
CCGTAATGTG GTAGACGGTC AGCCTTTCTC AAACTGGTGG GACAACGGGA GCAATCAAGT
AGCTTTCGGT CGCGGCGACA GAGGCTTCAT TGTCTTTAAT AATGATGACT GGTATATGAA
TGTCGATTTG CAAACTGGTC TGCCTGCTGG AACCTACTGC GATGTTATTT CTGGACAAAA
GGAAGGCAGT GCGTGTACTG GAAAGCAGGT GTACGTTTCC TCGGATGGAA AGGCCAATTT
CCAGATTAGT AACAGCGATG AAGATCCATT TGTTGCAATT CACGTTGATG CCAAGTTATA
AGCTT-3'

Figure 1

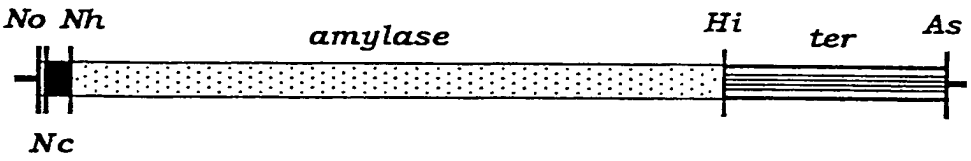

Figure 2

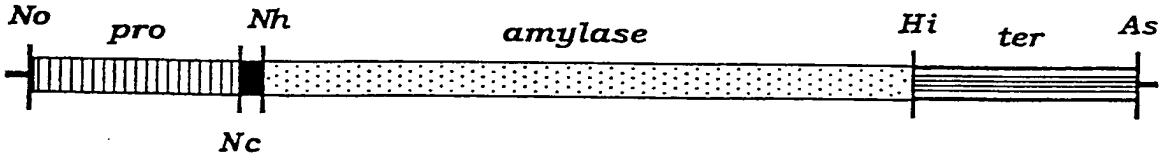

Figure 3

5'-AAGCTTGTAA ACAGCTGGGG AGCATGGCGA ACAGCCAGGC AATTAATTGA GATTATTAAT
TGTACGAAAT ATATATGATG AGATTATAAA CACACAACAC TTTTATTCGC AAGGGATGAT
AAGAATCTAA TATATATATT ATCTGGGCTT CAAAGCATTG ATTTTATTTA TTGAGTCAAG
AGGGAAATTT ATTTTCTTGT TATTCTCTGT CCAGGTCTAA AGTCCCGAGC GGTGAGGCTA
TCTATTGATT TGGACATTCC AATCGAATAC AAAACAGAGA TACAGAAATT TGCGAAAAAA
TTTGATAACA ATCGTGGATT TTACGAATTA GACAAATTGA TATGTGCTTG CTAATTGATG
TGGCATGAAA TAAGAAATTT ATAAGGACGT TTTCAAGTGC TTCTATTTTA AACATTCAGG
ATTTTTTTTT AAAGCAGACA GCTTTCAACA GGTTTGATGA GAATTTGAAT ATTGATTGTT
GACTTTAGCT ATACATAAAT CACACCTCAT CCACCCATTG TGGTATCCTT CGAAGGACTT
GGGAACTGGA TCCTCTAGAA GGCGCGCC-3'

Figure 4

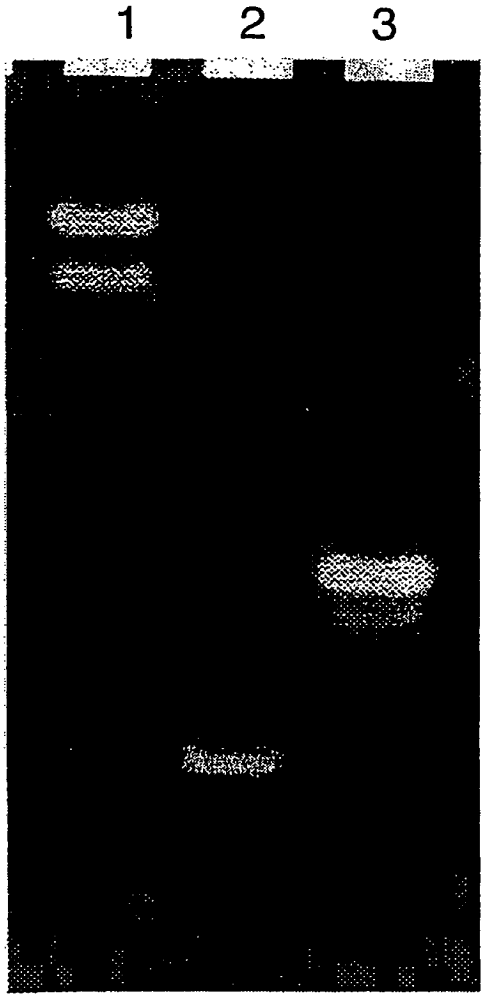

FIG. 5

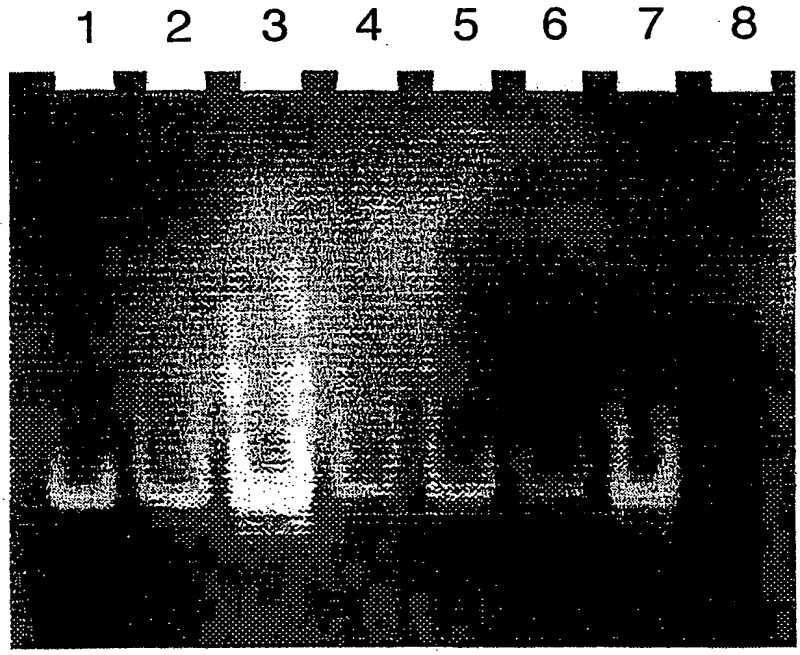

FIG. 6

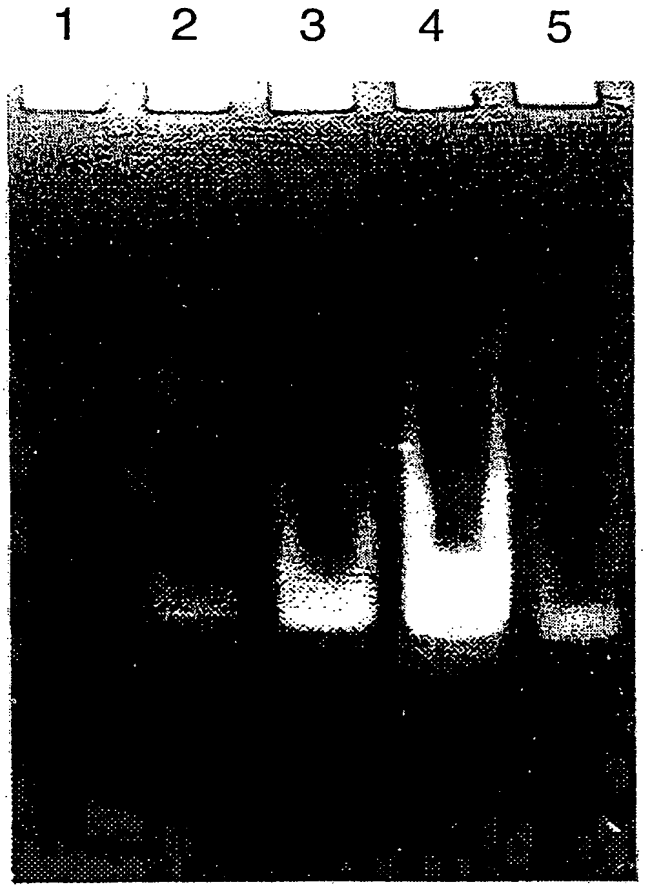

FIG. 7

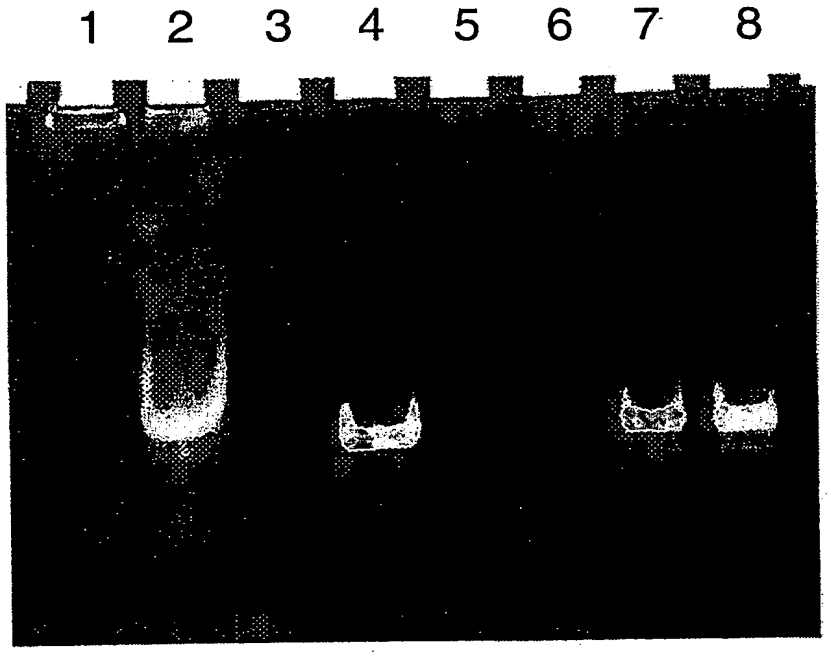

FIG. 8

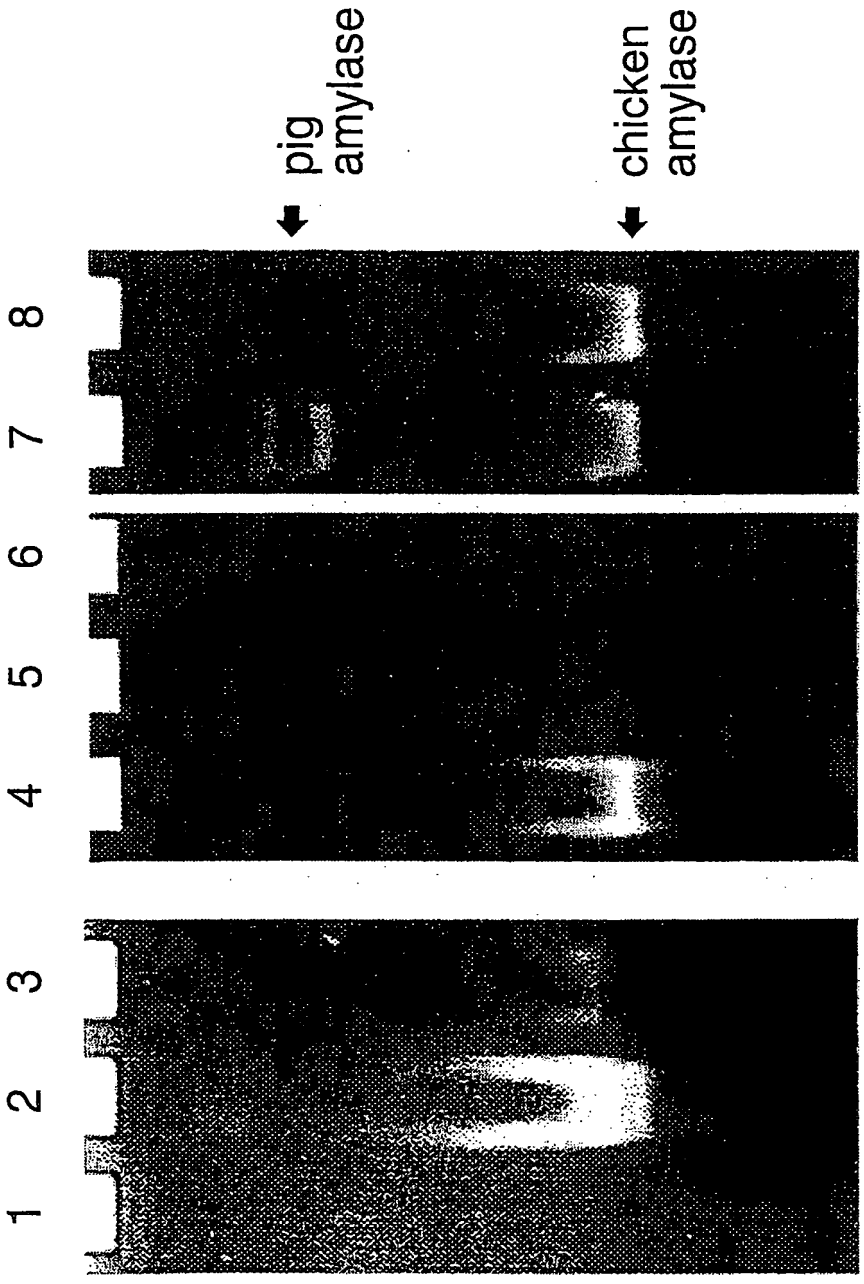

FIG. 9

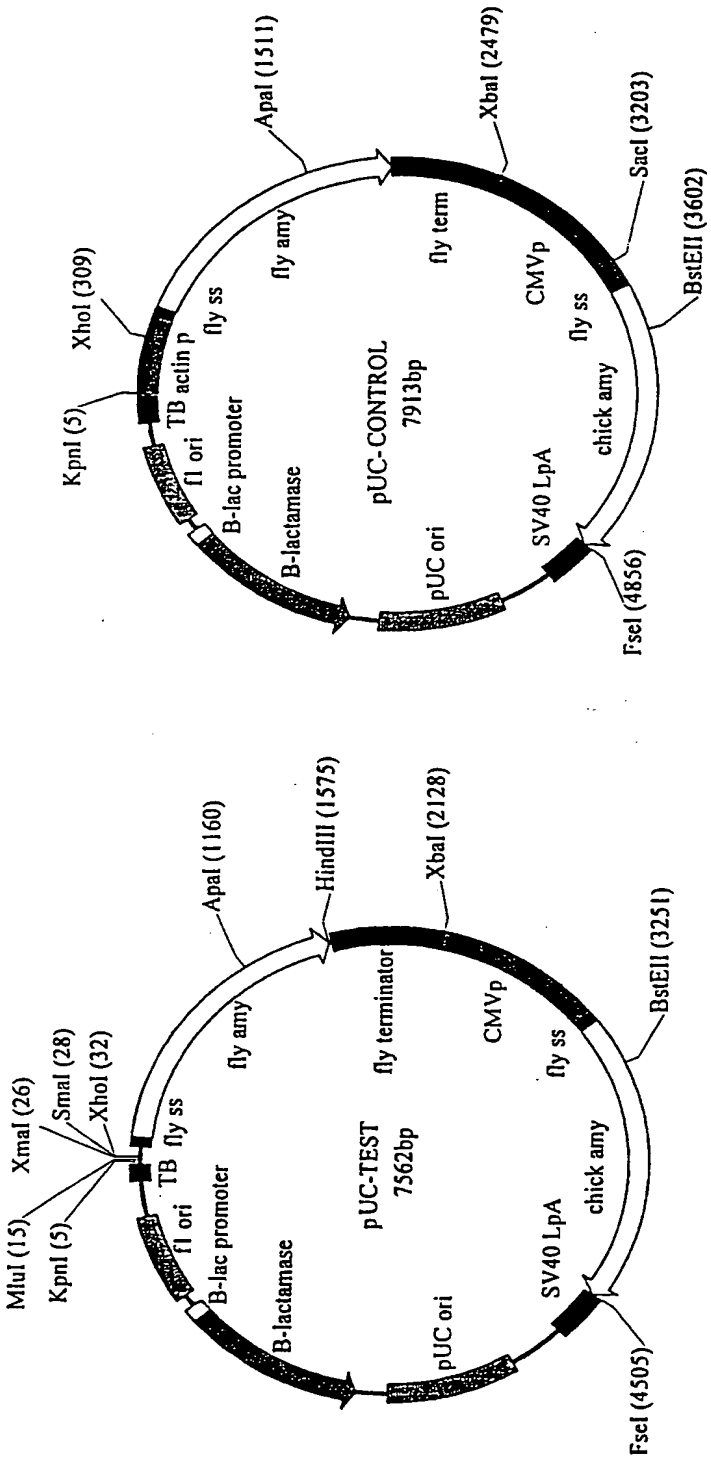
ApaI (1511)
XbaI (2479)
SacI (3203)
BstEII (3602)
XhoI (309)
KpnI (5)
TB actin p
fly ss
fly amy
fly term
CMVp
fly ss
chick amy
f1 ori
B-lac promoter
B-lactamase
pUC-CONTROL
7913bp
pUC ori
SV40 LpA
FseI (4856)
ApaI (1160)
HindIII (1575)
XbaI (2128)
BstEII (3251)
XmaI (26)
SmaI (28)
XhoI (32)
MluI (15)
KpnI (5)
TB fly ss
f1 ori
B-lac promoter
fly amy
fly terminator
CMVp
fly ss
chick amy
B-lactamase
pUC-TEST
7562bp
pUC ori
SV40 LpA
FseI (4505)

Figure 10

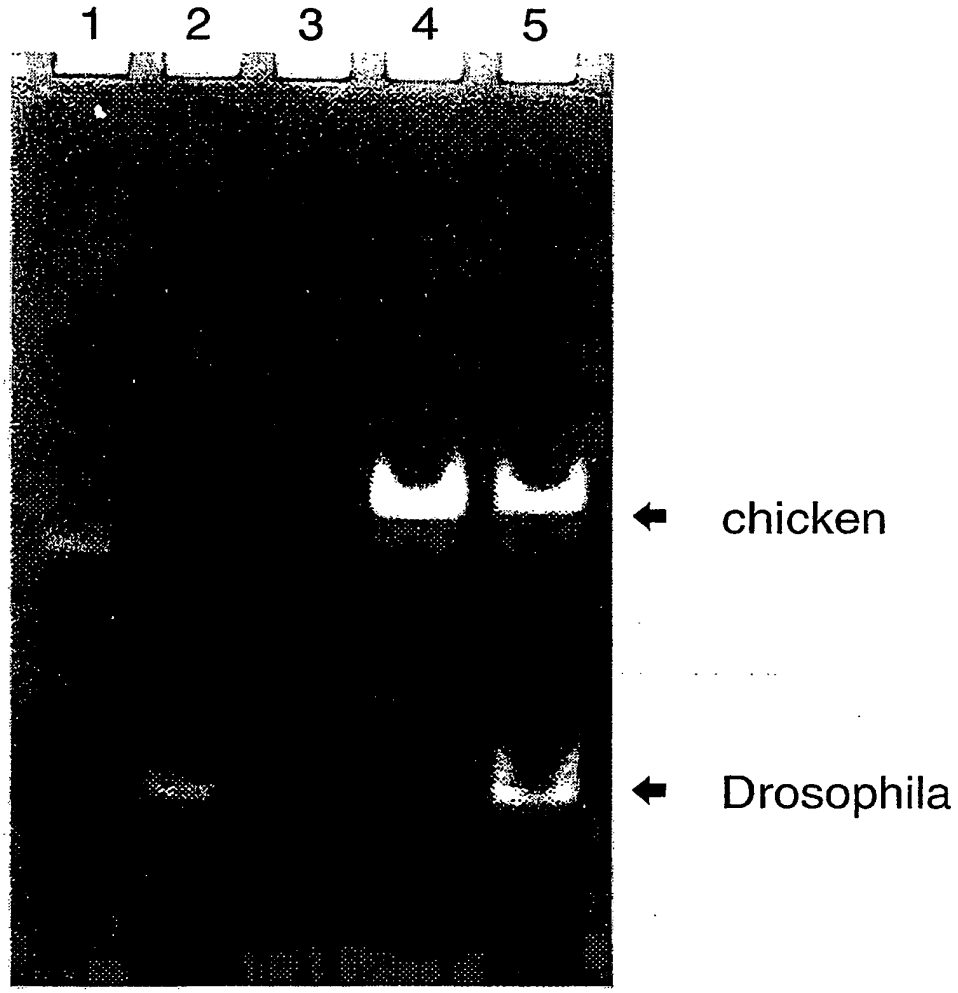

FIG. 11

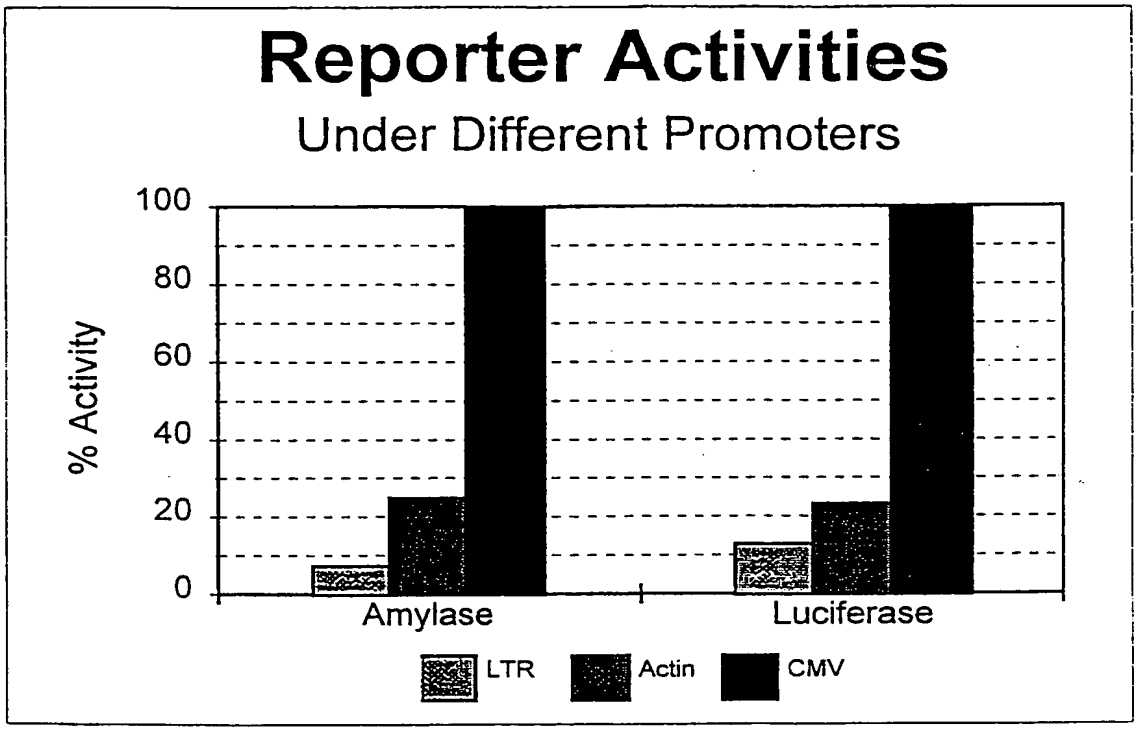
Reporter Activities
Under Different Promoters
% Activity
100
80
60
40
20
0
Amylase
Luciferase
LTR
Actin
CMV

Figure 12

INTERNATIONAL SEARCH REPORT

International Application No
PCT/CA 98/00157

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6 C12N15/56 C12N15/85 C12Q1/68

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 6 C12N C12Q

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	EP 0 438 200 A (CIGB) 24 July 1991 see the whole document ---	1,2,12, 30,32
A	BOER, P.H. AND HICKEY, D.A.: "The alpha-amylase gene in Drosophila melanogaster:nucleotide sequence,gene structure and expression motifs" NUCLEIC ACIDS RESEARCH, vol. 14, no. 21, 1986, pages 8399-8411, XP002068035 see the whole document --- -/--	1

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
15 June 1998	06/07/1998

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Hillenbrand, G

1

Form PCT/ISA/210 (second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

International Application No
PCT/CA 98/00157

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	INOMATA, N. ET AL.: "Evolutionary relationships and sequence variation of alpha-amylase variants encoded by duplicated genes in the amy locus of Drosophila melanogaster" GENETICS, vol. 141, 1995, pages 237-244, XP002068036 see the whole document	1
A	WISE, R.J. ET AL.: "A complementary DNA sequence that predicts a human pancreatic amylase primary structure consistent with the electrophoretic mobility of the common isozyme, amy2 A." MOL. BIOL. MED., vol. 2, 1984, pages 307-322, XP002068037 see the whole document	1

1

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No
PCT/CA 98/00157

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
EP 0438200 A	24-07-1991	JP 5184352 A	27-07-1993